

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

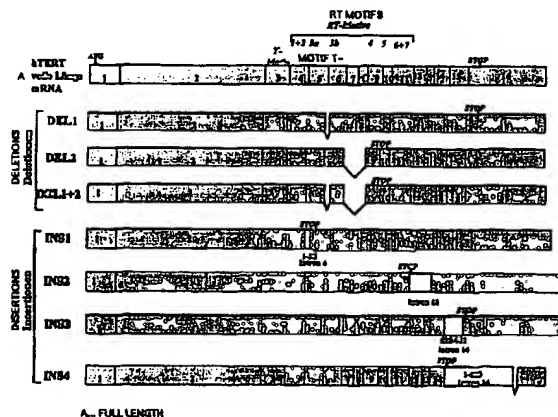


  
**PCT** WELTORGANISATION FÜR GEISTIGES EIGENTUM  
 Internationales Büro  
 INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE  
 INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

<p>(51) Internationale Patentklassifikation 6 :  <b>C12N 15/54, 9/12, 15/11, 15/85, C12Q 1/68, A01K 67/027</b></p>	<b>A2</b>	<p>(11) Internationale Veröffentlichungsnummer: <b>WO 99/33998</b></p> <p>(43) Internationales Veröffentlichungsdatum: 8. Juli 1999 (08.07.99)</p>		
<table style="width: 100%;"> <tr> <td style="width: 50%; vertical-align: top;"> <p>(21) Internationales Aktenzeichen: PCT/EP98/08216</p> <p>(22) Internationales Anmeldedatum: 22. Dezember 1998 (22.12.98)</p> <p>(30) Prioritätsdaten:            197 57 984.1      24. Dezember 1997 (24.12.97)    DE</p> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): BAYER AKTIENGESELLSCHAFT [DE/DE]; D-51368 Leverkusen (DE).</p> <p>(72) Erfinder; und            (75) Erfinder/Anmelder (nur für US): HAGEN, Gustav [DE/DE]; Bertha-von-Suttner-Strasse 31, D-51373 Leverkusen (DE). WICK, Maresa [DE/DE]; Andreas-Gryphius-Strasse 26, D-51065 Köln (DE). ZUBOV, Dmitry [RU/DE]; Roggen-dorfstrasse 59, D-51061 Köln (DE).</p> <p>(74) Gemeinsamer Vertreter: BAYER AKTIENGESELLSCHAFT; D-51368 Leverkusen (DE).</p> </td> <td style="width: 50%; vertical-align: top;"> <p>(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Veröffentlicht</b>  <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i></p> </td> </tr> </table>			<p>(21) Internationales Aktenzeichen: PCT/EP98/08216</p> <p>(22) Internationales Anmeldedatum: 22. Dezember 1998 (22.12.98)</p> <p>(30) Prioritätsdaten:            197 57 984.1      24. Dezember 1997 (24.12.97)    DE</p> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): BAYER AKTIENGESELLSCHAFT [DE/DE]; D-51368 Leverkusen (DE).</p> <p>(72) Erfinder; und            (75) Erfinder/Anmelder (nur für US): HAGEN, Gustav [DE/DE]; Bertha-von-Suttner-Strasse 31, D-51373 Leverkusen (DE). WICK, Maresa [DE/DE]; Andreas-Gryphius-Strasse 26, D-51065 Köln (DE). ZUBOV, Dmitry [RU/DE]; Roggen-dorfstrasse 59, D-51061 Köln (DE).</p> <p>(74) Gemeinsamer Vertreter: BAYER AKTIENGESELLSCHAFT; D-51368 Leverkusen (DE).</p>	<p>(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Veröffentlicht</b>  <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i></p>
<p>(21) Internationales Aktenzeichen: PCT/EP98/08216</p> <p>(22) Internationales Anmeldedatum: 22. Dezember 1998 (22.12.98)</p> <p>(30) Prioritätsdaten:            197 57 984.1      24. Dezember 1997 (24.12.97)    DE</p> <p>(71) Anmelder (für alle Bestimmungsstaaten ausser US): BAYER AKTIENGESELLSCHAFT [DE/DE]; D-51368 Leverkusen (DE).</p> <p>(72) Erfinder; und            (75) Erfinder/Anmelder (nur für US): HAGEN, Gustav [DE/DE]; Bertha-von-Suttner-Strasse 31, D-51373 Leverkusen (DE). WICK, Maresa [DE/DE]; Andreas-Gryphius-Strasse 26, D-51065 Köln (DE). ZUBOV, Dmitry [RU/DE]; Roggen-dorfstrasse 59, D-51061 Köln (DE).</p> <p>(74) Gemeinsamer Vertreter: BAYER AKTIENGESELLSCHAFT; D-51368 Leverkusen (DE).</p>	<p>(81) Bestimmungsstaaten: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Veröffentlicht</b>  <i>Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.</i></p>			

(54) Title: REGULATORY DNA SEQUENCES OF THE HUMAN CATALYTIC TELOMERASE SUB-UNIT GENE, DIAGNOSTIC AND THERAPEUTIC USE THEREOF

(54) Bezeichnung: REGULATORISCHE DNA-SEQUENZEN DES GENS DER HUMANEN KATALYTISCHEN TELOMERASE-UNTEREINHEIT UND DEREN DIAGNOSTISCHE UND THERAPEUTISCHE VERWENDUNG



(57) Abstract

The present invention relates to regulatory DNA sequences containing promotor sequences, in addition to intervening sequences, for the human catalytic telomerase sub-unit gene. The invention also relates to the use of said DNA sequences for pharmaceutical, diagnostic and therapeutic purposes, especially in the treatment of cancer and ageing.

(57) Zusammenfassung

Diese Erfindung betrifft regulatorische DNA-Sequenzen, beinhaltend Promotorsequenzen, sowie Intronsequenzen, für das Gen der humanen kalytischen Telomerase-Untereinheit. Darüber hinaus betrifft diese Erfindung die Verwendung dieser DNA-Sequenzen für pharmazeutische, diagnostische und therapeutische Zwecke, vor allem in der Behandlung von Krebs und Alterung.

**LEDIGLICH ZUR INFORMATION**

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische Republik Mazedonien	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland	ML	Mali	TR	Türkei
BG	Bulgarien	HU	Ungarn	MN	Mongolei	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MR	Mauritanien	UA	Ukraine
BR	Brasilien	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Island	MX	Mexiko	US	Vereinigte Staaten von Amerika
CA	Kanada	IT	Italien	NE	Niger	UZ	Usbekistan
CF	Zentralafrikanische Republik	JP	Japan	NL	Niederlande	VN	Vietnam
CG	Kongo	KE	Kenia	NO	Norwegen	YU	Jugoslawien
CH	Schweiz	KG	Kirgisistan	NZ	Neuseeland	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik Korea	PL	Polen		
CM	Kamerun	KR	Republik Korea	PT	Portugal		
CN	China	KZ	Kasachstan	RO	Rumänien		
CU	Kuba	LC	St. Lucia	RU	Russische Föderation		
CZ	Tschechische Republik	LI	Liechtenstein	SD	Sudan		
DE	Deutschland	LK	Sri Lanka	SE	Schweden		
DK	Dänemark	LR	Liberia	SG	Singapur		
EE	Estland						

Regulatorische DNA-Sequenzen des Gens der humanen katalytischen  
Telomerase-Untereinheit und deren diagnostische und therapeutische Verwen-  
dung

5 Aufbau und Funktion der Chromosomenenden

Das genetische Material eukaryontischer Zellen ist auf linearen Chromosomen verteilt. Die Enden der Erbanlagen werden, abgeleitet von den griechischen Wörtern *telos* (Ende) und *meros* (Teil, Segment), als Telomere bezeichnet. Die meisten  
10 Telomere bestehen aus Wiederholungen von kurzen Sequenzen, die überwiegend aus Thymin und Guanin aufgebaut sind (Zakian, 1995). In allen bislang untersuchten Wirbeltieren werden die Telomere aus der Sequenz TTAGGG aufgebaut (Meyne *et al.*, 1989).

15 Die Telomere üben verschiedene wichtige Funktionen aus. Sie verhindern die Fusion von Chromosomen (McClintock, 1941) und damit die Entstehung von dizentrischen Erbanlagen. Solche Chromosomen mit zwei Centromeren können durch Verlust der Heterozygotie bzw. Verdopplung oder Verlust von Genen zur Entwicklung von Krebs führen.

20 Desweiteren dienen Telomere dazu, intakte Erbanlagen von beschädigten zu unterscheiden. So stellten Hefezellen ihre Zellteilung ein, wenn sie ein Chromosom ohne Telomer enthielten (Sandell und Zakian, 1993).

25 Eine weitere wichtige Aufgabe erfüllen Telomere bei der DNA-Replikation eukaryontischer Zellen. Im Gegensatz zu den zirkulären Genomen von Prokaryonten können die linearen Chromosomen der Eukaryonten von dem DNA Polymerase-Komplex nicht vollständig repliziert werden. Zur Initiation der DNA-Replikation sind RNA-Primer notwendig. Nach Abspaltung der RNA-Primer, Verlängerung der  
30 Okazaki-Fragmente und anschließender Ligation fehlt dem neu-synthetisierten DNA-Strang das 5'-Ende, denn dort kann der RNA-Primer nicht durch DNA ersetzt

werden. Ohne besondere Schutzmechanismen würden daher die Chromosomen mit jeder Zellteilung schrumpfen ("end-replication problem"; Harley *et al.*, 1990). Die nicht-kodierenden Telomersequenzen stellen vermutlich eine Pufferzone dar, um dem Verlust von Genen vorzubeugen (Sandell und Zakian, 1993).

5

Darüberhinaus spielen Telomere auch eine wichtige Rolle bei der Regulation der zellulären Alterung (Olovnikov, 1973). Humane somatische Zellen zeigen in Kultur eine limitierte Replikationskapazität; sie werden nach einer gewissen Zeit seneszent. In diesem Zustand teilen sich die Zellen selbst nach Stimulierung mit Wachstumsfaktoren nicht mehr, sterben aber nicht, sondern bleiben metabolisch aktiv (Goldstein, 1990). Verschiedene Beobachtungen sprechen für die Hypothese, daß eine Zelle anhand der Länge ihrer Telomere bestimmt, wie oft sie sich noch teilen kann (Allsopp *et al.*, 1992).

10

15 Zusammenfassend besitzen die Telomere somit zentrale Funktionen bei der Alterung von Zellen sowie der Stabilisierung des genetischen Materials und Verhinderung von Krebs.

#### Das Enzym Telomerase synthetisiert die Telomere

20

Wie oben beschrieben können Organismen mit linearen Chromosomen ohne einen speziellen Schutzmechanismus ihr Genom nur unvollständig replizieren. Die meisten Eukaryonten verwenden zur Regeneration der Telomersequenzen ein spezielles Enzym, die Telomerase. In den bislang untersuchten Einzellern wird Telomerase konstitutiv exprimiert. Dagegen wurde in Menschen die Telomerase-Aktivität nur in Keimzellen und Tumorzellen gemessen, wogegen benachbartes somatisches Gewebe keine Telomerase enthielt (Kim *et al.*, 1994).

25

Funktionell kann die Telomerase auch als terminale Telomertransferase bezeichnet werden, die als Multiproteinkomplex im Zellkern lokalisiert ist. Während der RNA-Anteil der humanen Telomerase schon seit längerem bekannt ist (Feng *et al.*, 1995),

30

wurde kürzlich die katalytische Untereinheit dieser Enzymgruppe in verschiedenen Organismen identifiziert (Lingner *et al.*, 1997; vgl. unsere ebenfalls anhängige Anmeldung PCT EP/98/03468). Diese katalytischen Untereinheiten der Telomerase sind sowohl untereinander als auch zu bisher allen bekannten reversen Transkriptasen auffällig homolog.

Auch in WO 98/14592 werden Nukleinsäure- und Aminosäuresequenzen der katalytischen Telomerase-Untereinheit beschrieben.

#### 10 Aktivierung der Telomerase in menschlichen Tumoren

Eine Aktivität der Telomerase konnte in Menschen ursprünglich nur in Keimbahnzellen, nicht aber in normalen somatischen Zellen (Hastie *et al.*, 1990; Kim *et al.*, 1994) nachgewiesen werden. Nach der Entwicklung eines sensitiveren Nachweisverfahrens (Kim *et al.*, 1994) wurde auch in hematopoietischen Zellen eine geringe Telomeraseaktivität detektiert (Broccoli *et al.*, 1995; Counter *et al.*, 1995; Hiyama *et al.*, 1995). Allerdings wiesen diese Zellen trotzdem eine Reduktion der Telomere auf (Vaziri *et al.*, 1994; Counter *et al.*, 1995). Noch ist nicht geklärt, ob die Menge an Enzym in diesen Zellen nicht ausreichend für eine Kompensation des Telomerverlustes ist, oder ob die gemessene Telomerase-Aktivität von einer Subpopulation, z.B. unvollständig ausdifferenzierten CD34<sup>+</sup>38<sup>+</sup>-Vorläuferzellen, herrührt (Hiyama *et al.*, 1995). Zur Klärung wäre ein Nachweis der Telomerase-Aktivität in einer einzelnen Zelle nötig.

Interessanterweise wurde jedoch in einer großen Zahl der bislang getesteten Tumorgewebe eine signifikante Telomerase-Aktivität nachgewiesen (1734/2031, 85 %; Shay, 1997), während in normalem somatischem Gewebe keine Aktivität gefunden wurde (1/196, <1 %, Shay, 1997). Verschiedene Untersuchungen zeigten außerdem, daß in seneszenten Zellen, die mit viralen Oncoproteinen transformiert wurden, die Telomere weiterhin schrumpften und Telomerase nur in der Subpopulation entdeckt werden konnte, die die Wachstumskrise überlebte (Counter *et al.*, 1992). In diesen immortalisierten Zellen waren auch die Telomere stabil (Counter *et al.*, 1992). Ähnli-

che Befunde aus Untersuchungen an Mäusen (Blasco *et al.*, 1996) stützen die Annahme, daß eine Reaktivierung der Telomerase ein spätes Ereignis in der Tumorgenese ist.

5       Basierend auf diesen Ergebnissen wurde eine "Telomerase-Hypothese" entwickelt, die den Verlust von Telomersequenzen und Zellalterung mit der Aktivität von Telomerase und der Entstehung von Krebs verbindet. In langlebigen Spezies wie dem Menschen kann das Schrumpfen der Telomere als ein Mechanismus zur Tumorsuppression angesehen werden. Ausdifferenzierte Zellen, die keine Telomerase  
10       enthalten, stellen bei einer bestimmten Länge der Telomere ihre Zellteilung ein. Mutiert eine solche Zelle, so kann aus ihr nur dann ein Tumor entstehen, wenn die Zelle ihre Telomere verlängern kann. Ansonsten würde die Zelle weiterhin Telomersequenzen verlieren, bis ihre Chromosomen instabil werden und sie schließlich zugrunde geht. Die Reaktivierung der Telomerase ist vermutlich der Hauptmechanismus  
15       von Tumorzellen zur Stabilisation ihrer Telomere.

Aus diesen Beobachtungen und Überlegungen ergibt sich, daß eine Inhibition der Telomerase eine Therapie von Tumoren erlauben sollte. Konventionelle Krebstherapien mit Zytostatika oder kurzweiligen Strahlen schädigen nicht nur die Tumorzellen,  
20       sondern alle sich teilenden Zellen des Körpers. Da aber außer Tumorzellen nur Keimbahnzellen eine signifikante Telomerase-Aktivität enthalten, würden Telomerase-Inhibitoren spezifischer die Tumorzellen angreifen und somit weniger unerwünschte Nebenwirkungen hervorrufen. In allen bislang getesteten Tumorgeweben wurde eine Telomerase-Aktivität nachgewiesen, so daß diese Therapeutika gegen alle  
25       Krebsarten eingesetzt werden könnten. Die Wirkung von Telomerase-Inhibitoren würde dann eintreten, wenn die Telomere der Zellen sich soweit verkürzt haben, daß das Genom instabil wird. Da Tumorzellen meist kürzere Telomere aufweisen als normale somatische Zellen, würden zuerst Krebszellen durch Telomerase-Inhibitoren eliminiert werden. Zellen mit langen Telomeren, wie die Keimzellen, würden  
30       dagegen erst viel später geschädigt werden. Telomerase-Inhibitoren stellen somit einen zukunftsweisenden Weg für die Therapie von Krebs dar.

Eindeutige Antworten auf die Frage nach der Art und den Angriffspunkten physiologischer Telomerase-Inhibitoren werden möglich sein, wenn auch die Regulation der Genexpression der Telomerase identifiziert ist.

5

#### Regulation der Genexpression in Eukaryonten

Die eukaryotische Genexpression, d.h. der zelluläre Informationsfluß von der DNA über die RNA zum Protein, weist vielfältige Ansatzpunkte für regulatorische Mechanismen auf. Einzelne Kontrollstufen sind z.B. die Gen-Amplifikation, Rekombination von Genloci, Chromatinstruktur, DNA-Methylierung, Transkription, posttranskriptionelle mRNA-Modifikationen, mRNA-Transport, Translation und post-translationale Proteinmodifikationen. Nach bisherigen Studien besitzt die Kontrolle auf der Ebene der Transkriptionsinitiation die größte Bedeutung (Latchman, 1991).

15

Unmittelbar stromaufwärts vom Transkriptionsstart eines von der RNA-Polymerase II transkribierten Gens liegt eine Region, die für die Steuerung der Transkription verantwortlich ist und als Promotorregion bezeichnet wird. Ein Vergleich der Nukleotidsequenzen von Promotorregionen vieler bekannter Gene zeigt, daß bestimmte Sequenzmotive in dieser Region häufig vorkommen. Zu diesen Elementen gehören unter anderem die TATA-Box, die CCAAT-Box und die GC-Box, die von spezifischen Proteinen erkannt werden. Die TATA-Box, die etwa 30 Nukleotide stromaufwärts vom Transkriptionsstart entfernt positioniert ist, wird z.B. von der TFIID-Untereinheit TBP („TATA-box binding protein“) erkannt, wogegen bestimmte GC-reiche Sequenzabschnitte vom Transkriptionsfaktor Sp1 („specificity protein1“) spezifisch gebunden werden.

20

25

Funktionell kann man den Promotor in einen regulativen und einen konstitutiven Abschnitt unterteilen (Latchman, 1991). Der konstitutive Kontrollbereich umfaßt den sogenannten Kernpromotor („corepromoter“), der die korrekte Initiation der Transkription ermöglicht. Er enthält die als UPE's (upstream promoter elements“) be-

30

schriebenen Sequenzelemente, die für eine effiziente Transkription notwendig sind. Die regulativen Kontrollabschnitte, die mit den UPE's verflochten sein können, weisen Sequenzelemente auf, die an der signalabhängigen Regulation der Transkription durch Hormone, Wachstumsfaktoren usw. beteiligt sein können. Sie vermitteln gewebs- oder zellspezifische Promotoreigenschaften.

Ein charakteristisches Merkmal eukaryotischer Gene sind DNA-Abschnitte, die über vergleichsweise große Distanzen hinweg Einfluß auf die Genexpression nehmen können. Diese Elemente können stromaufwärts, stromabwärts oder innerhalb einer Transkriptionseinheit lokalisiert sein und unabhängig von ihrer Orientierung ihre Funktion wahrnehmen. Diese Sequenzabschnitte können die Promotoraktivität verstärken (Enhancer) oder abschwächen (Silencer). Ähnlich wie die Promotorregionen beherbergen auch Enhancer und Silencer mehrere Bindungsstellen für Transkriptionsfaktoren.

Die Erfindung betrifft die DNA-Sequenzen aus der 5'-flankierenden Region des Gens der katalytisch aktiven humanen Telomerase-Untereinheit sowie Intron-Sequenzen für dieses Gen.

Die Erfindung betrifft insbesondere die 5'-flankierende regulatorische DNA-Sequenz, enthaltend die Promotor-DNA-Sequenz für das Gen der humanen katalytischen Telomerase Untereinheit gemäß Fig. 10 (SEQ ID NO 3).

Die Erfindung betrifft weiterhin regulatorisch wirksame Teilbereiche der 5'-flankierenden regulatorischen DNA-Sequenz gemäß Fig. 4 (SEQ ID NO 1).

Weiterhin sind Gegenstand der vorliegenden Erfindung Intron-Sequenzen für das Gen der humanen katalytischen Telomerase-Untereinheit, insbesondere solche, die regulatorische Wirkung haben. Die erfindungsgemäßen Intronsequenzen werden im Rahmen von Beispiel 5 detailliert beschrieben (vgl. SEQ ID NO 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 und 20).

Die Erfindung betrifft weiterhin ein rekombinantes Konstrukt, das die erfindungsgemäßen DNA-Sequenzen, insbesondere die 5'-flankierende DNA-Sequenz des Gens der humanen katalytischen Telomerase Untereinheit oder Teilbereiche davon  
5 beinhaltet.

Bevorzugt sind rekombinante Konstrukte, die neben den erfindungsgemäßen DNA-Sequenzen, insbesondere der 5'-flankierenden DNA-Sequenz des Gens der humanen katalytischen Telomerase Untereinheit oder Teilbereichen davon, eine oder mehrere  
10 weitere DNA-Sequenzen, die für Polypeptide oder Proteine kodieren, enthalten.

Gemäß einer besonders bevorzugten Ausführungsform kodieren diese weiteren DNA-Sequenzen für antitumorale Proteine.

15 Besonders bevorzugte antitumorale Proteine sind solche, die die Angiogenese direkt oder indirekt inhibieren. Zu diesen Proteinen zählen beispielsweise:

Plasminogenaktivatorinhibitor (PAI-1), PAI-2, PAI-3, Angiostatin, Endostatin, Platelet factor 4, TIMP-1, TIMP-2, TIMP-3, Leukemia Inhibitory Factor (LIF).  
20

Ebenfalls besonders bevorzugt sind antitumorale Proteine, welche direkt oder indirekt eine zytostatische Wirkung auf Tumoren aufweisen. Hierzu zählen im besonderen:

25 Perforin, Granzym, IL-2, IL-4, IL-12, Interferone, wie beispielsweise IFN- $\alpha$ , IFN- $\beta$ , IFN- $\gamma$ , TNF, TNF- $\alpha$ , TNF- $\beta$ , Oncostatin M; Tumorsuppressorgene, wie z.B. p53, Retinoblastoma.

Weiterhin besonders bevorzugt sind antitumorale Proteine, welche gegebenenfalls  
30 zusätzlich zur antitumoralen Wirkung Entzündungen stimulieren und hierdurch zur Elimination von Tumorzellen beitragen. Hierzu zählen beispielsweise:

RANTES, Monocyte chemotactic and activating factor (MCAF), IL-8, Macrophage inflammatory protein (MIP-1 $\alpha$ , - $\beta$ ), Neutrophil activating protein-2 (NAP-2), IL-3, IL-5, human leukemia inhibitory factor (LIF), IL-7, IL-11, IL-13, GM-CSF, G-CSF, M-CSF.

Weiterhin besonders bevorzugt sind antitumorale Proteine, welche aufgrund ihrer Wirkung als Enzyme in der Lage sind, Vorstufen eines antitumoralen Wirkstoffes in einen antitumoralen Wirkstoff zu überführen. Zu diesen Enzymen zählen beispielsweise:

Herpes Simplex Virus Thymidinkinase, Varizella Zoster Virus Thymidinkinase, bakterielle Nitroreductase, bakterielle  $\beta$ -Glukuronidase, pflanzliche  $\beta$ -Glukuronidase aus *Secale cereale*, humane Glukuronidase, humane Carboxypeptidase, bakterielle Carboxypeptidase, bakterielle  $\beta$ -Lactamase, bakterielle Cytosindeaminidase, humane Katalase bzw. Phosphatase, humane alkalische Phosphatase, Typ 5 saure Phosphatase, humane Lysooxidase, humane saure D-Aminooxidase, humane Glutathion Peroxidase, humane Eosinophilen Peroxidase, humane Schilddrüsen Peroxidase.

Die obengenannten rekombinanten Konstrukte können auch DNA-Sequenzen enthalten, die für Faktor VIII, IX oder Teilfragmente davon kodieren. Zu diesen DNA-Sequenzen zählen auch andere Blutgerinnungsfaktoren

Die obengenannten rekombinanten Konstrukte können auch DNA-Sequenzen enthalten, die für ein Reporterprotein kodieren. Zu diesen Reporterproteinen zählen beispielsweise:

Chloramphenicolacetyltransferase (CAT), Glühwürmchen Luziferase (LUC),  $\beta$ -Galaktosidase ( $\beta$ -Gal), Sezernierte alkalische Phosphatase (SEAP), Humanes Wachstumshormon (hGH),  $\beta$ -Glukuronidase (GUS), Grün-fluoreszierendes Protein (GFP) und alle davon abgeleiteten Varianten, Aquarin, Obelin.

Erfindungsgemäße rekombinante Konstrukte können auch DNA kodierend für die humane katalytische Telomerase Untereinheit und deren Varianten und Fragmente in antisense Orientierung enthalten. Gegebenenfalls können diese Konstrukte auch  
5 andere Protein-Untereinheiten der humanen Telomerase und die Telomerase-RNA-Komponente in antisense Orientierung enthalten.

Die rekombinanten Konstrukte können neben der DNA, kodierend für die humane katalytische Telomerase Untereinheit, sowie deren Varianten und Fragmente auch  
10 andere Protein-Untereinheiten der humanen Telomerase und die Telomerase-RNA-Komponente enthalten.

Die Erfindung betrifft weiterhin einen Vektor, enthaltend die oben genannten erfindungsgemäßen DNA-Sequenzen, insbesondere die 5'-flankierenden DNA-Sequenzen, sowie eine oder mehrere der oben genannten anderen DNA-Sequenzen.  
15

Bevorzugter Vektor für solche Konstrukte ist ein Virus, beispielsweise ein Retrovirus, Adenovirus, adeno-assoziiertes Virus, Herpes Simplex Virus, Vaccina Virus, lentivirales Virus, Sindbis Virus und ein Semliki Forest Virus.  
20

Ebenfalls bevorzugt sind Plasmide als Vektoren.

Die Erfindung betrifft weiterhin pharmazeutische Präparate, enthaltend erfindungsgemäße rekombinante Konstrukte bzw. Vektoren; beispielsweise eine Zubereitung in einem kolloidalen Dispersionssystem.  
25

Geeignete kolloidale Dispersionssysteme sind beispielsweise Liposome oder Polylysin-Liganden.

Die Zubereitungen der erfindungsgemäßen Konstrukte bzw. Vektoren in kolloidalen Dispersionssystemen können um einen Liganden ergänzt sein, der an Membranstruk-  
30

turen von Tumorzellen bindet. Ein solcher Ligand kann z.B. an das Konstrukt bzw. den Vektor angeknüpft sein oder auch Bestandteil der Liposomenstruktur sein.

5 Geeignete Liganden sind insbesondere polyklonale oder monoklonale Antikörper oder Antikörperfragmente hiervon, die mit ihren variablen Domänen an Membranstrukturen von Tumorzellen binden, oder endständige Mannose-tragende Substanzen, Zytokine, Wachstumsfaktoren oder Fragmente bzw. Teilsequenzen hiervon, die an Rezeptoren auf Tumorzellen binden.

10 Entsprechende Membranstrukturen sind beispielsweise Rezeptoren für ein Zytokin oder einen Wachstumsfaktor, wie z.B. IL-1, EGF, PDGF, VEGF, TGF  $\beta$ , Insulin oder Insulin-like Growth Factor (IGF), oder Adhäsionsmoleküle, wie z. B. SLeX, LFA-1, MAC-1, LECAM-1 oder VLA-4, oder der Mannose-6-Phosphat-Rezeptor.

15 Zur vorliegenden Erfindung gehören pharmazeutische Zubereitungen, die neben den erfindungsgemäßen Vektorkonstrukten auch nichttoxische, inerte, pharmazeutisch geeignete Trägerstoffe enthalten können. Vorstellbar sind die Applikation (z.B. intravenös, intraarteriell, intramuskulär, subkutan, intradermal, anal, vaginal, nasal, transdermal, intraperitoneal, als Aerosol oder oral) am Ort eines Tumors oder die systemische Applikation dieser Zubereitungen.

20

Die erfindungsgemäßen Vektorkonstrukte können in der Gentherapie eingesetzt werden.

25 Die Erfindung betrifft weiterhin eine rekombinante Wirtszelle, insbesondere eine rekombinante eukaryotische Wirtszelle, enthaltend die vorstehend beschriebenen Konstrukte bzw. Vektoren.

30 Die Erfindung betrifft weiterhin ein Verfahren zur Identifizierung von Substanzen, die die Promotor-, Silencer- oder Enhanceraktivität der katalytischen Telomerase Untereinheit beeinflussen, wobei dieses Verfahren folgende Schritte umfaßt:

A. Zugabe einer Kandidatensubstanz zu einer Wirtszelle, enthaltend die erfindungsgemäße regulatorische DNA-Sequenz, insbesondere die 5'-flankierende regulatorische DNA-Sequenz für das Gen der humanen katalytischen  
5 Telomerase-Untereinheit oder einen regulatorisch wirksamen Teilbereich davon, funktionell verknüpft mit einem Reportergen,

B. Messung des Substanzeffektes auf die Reportergenexpression.

10 Das Verfahren kann eingesetzt werden zur Identifizierung von Substanzen, die die Promotor-, Silencer- oder Enhanceraktivität der katalytischen Telomerase Untereinheit verstärken.

Das Verfahren kann weiterhin eingesetzt werden zur Identifizierung von Substanzen,  
15 die die Promotor-, Silencer- oder Enhanceraktivität der katalytischen Telomerase Untereinheit inhibieren.

Die Erfindung betrifft weiterhin ein Verfahren zur Identifizierung von Faktoren, die spezifisch an Fragmente der erfindungsgemäßen DNA-Fragmente, insbesondere der  
20 5'-flankierenden regulatorischen DNA-Sequenz der katalytischen Telomerase Untereinheit, binden. Diese Methode beinhaltet ein Screening einer Expressions-cDNA-Bibliothek mit der vorstehend beschriebenen DNA-Sequenz oder Teilfragmenten unterschiedlichster Länge als Sonde.

25 Die vorstehend beschriebenen Konstrukte bzw. Vektoren können auch zur Herstellung transgener Tiere verwendet werden.

Die Erfindung betrifft weiterhin ein Verfahren zur Detektion Telomerase-assoziiierter Zustände bei einem Patienten, das folgende Schritte umfaßt:

30

- A. Inkubation eines Konstruktes bzw. Vektors, enthaltend die erfindungsgemäße DNA-Sequenz, insbesondere die 5'-flankierende regulatorische DNA-Sequenz für das Gen der humanen katalytischen Telomerase-Untereinheit oder einen regulatorisch wirksamen Teilbereich davon sowie ein Reportergen mit Körperflüssigkeiten oder zellulären Proben,
- B. Detektion der Reportergenaktivität, um einen diagnostischen Wert zu erhalten;
- C. Vergleich des diagnostischen Werts mit Standardwerten für das Reportergenkonstrukt in standardisierten normalen Zellen oder Körperflüssigkeiten des gleichen Typs wie die Testprobe;

Detektion diagnostischer Werte, die höher oder niedriger als Standardvergleichswerte liegen, indiziert einen Telomerase-assoziierten Zustand, der wiederum einen pathogenen Zustand indiziert.

Erläuterung der Abbildungen:

Fig. 1: Southern Blot-Analyse mit genomischer DNA verschiedener Spezies

A: Foto eines Ethidiumbromid gefärbten 0,7 %igen Agarosegels mit etwa 4 µg Eco RI geschnittener genomischer DNA. Die Spur 1 enthält Hind III geschnittene λ-DNA als Größenmarker (23,5, 9,4, 6,7, 4,4, 2,3, 2,0, und 0,6 kb). Die Spuren 2 bis 10 enthalten genomische DNA von Mensch, Rhesusaffe, Sprague Dawley Ratte, BALB/c Maus, Hund, Rind, Kaninchen, Huhn und Hefe (*Saccharomyces cerevisiae*).

B: Zu Fig.1 A korrespondierendes Autoradiogramm einer Southern Blot-Analyse, hybridisiert mit einer radioaktiv-markierten etwa 720 bp langen hTC-cDNA Sonde.

Fig. 2: Restriktionsanalyse der rekombinanten  $\lambda$ -DNA des Phagenklons P12, der mit einer Sonde aus dem 5'-Bereich der hTC-cDNA hybridisiert.

5 Die Abbildung zeigt ein Foto eines Ethidiumbromid gefärbten 0,4 %igen Agarosegels. Die Spuren 1 und 2 enthalten Eco RI/Hind III geschnittene  $\lambda$ -DNA bzw eine 1 kb Leiter der Firma Gibco als Größenmarker. Die Spuren 3 - 7 enthalten 250 ng mit Bam HI (Spur 3), Eco RI (Spur 4), Sal I (Spur 5), Xho I (Spur 6) und Sac I (Spur 7) geschnittene DNA des  
10 rekombinanten Phagens. Die Pfeile kennzeichnen die zwei  $\lambda$ -Arme des Vektors EMBL3 Sp6/T7.

Fig. 3: Restriktionsanalyse und Southern Blot-Analyse der rekombinanten  $\lambda$ -DNA des Phagenklons, der mit einer Sonde aus dem 5'-Bereich der hTC-cDNA hybridisiert.

15 A: Die Abbildung zeigt ein Foto eines Ethidiumbromid gefärbten 0,8%igen Agarosegels. Die Spuren 1 und 15 enthalten eine 1 kb Leiter der Firma Gibco als Größenmarker. Die Spuren 2 bis 14 enthalten 250 ng geschnittene  $\lambda$ -DNA vom rekombinanten Phagenklon. Als Enzyme wurden eingesetzt: Spur 2: Sac I, Spur 3: Xho I, Spur 4: Xho I, Xba I, Spur 5: Sac I, Xho I, Spur 6: Sal I, Xho I, Xba I, Spur 7: Sac I, Xho I, Xba I, Spur 8: Sac I, Sal I, Xba I, Spur 9: Sac I, Sal I, BamH I, Spur 10: Sac I, Sal I, Xho I, Spur 11: Not I, Spur 12: Sma I, Spur 13: leer, Spur  
20 14: nicht verdaut.

25 B: Zu Fig.3 A korrespondierendes Autoradiogramm einer Southern Blot-Analyse. Als Sonde für die Hybridisierung wurde ein etwa 420 bp langes 5'-hTC-cDNA Fragment eingesetzt.

30

Fig. 4: Partielle DNA-Sequenz der 5'-flankierenden Region und des Promotors vom Gen der humanen katalytischen Telomerase-Untereinheit. Das ATG-Startcodon ist in der Sequenz fett hervorgehoben. Die dargestellte Sequenz entspricht SEQ ID NO 1.

5

Fig. 5: Identifizierung des Transkriptionsstarts durch Primer Extension-Analyse.

10

Die Abbildung zeigt ein Autoradiogramm eines denaturierenden Polyacrylamidgels, welches zur Darstellung einer Primer Extension-Analyse gewählt wurde. Als Primer wurde ein Oligonukleotid mit der Sequenz 5'GTAAAGTTGTAGCTTAACTGGTTCTC 3' benutzt. In der Spur 1 wurde die Primer Extension Reaktion aufgetragen. Die Spuren G, A, T, C, stellen die Sequenzreaktionen mit dem gleichen Primer und den entsprechenden Dideoxynukleotiden dar. Der fette Pfeil kennzeichnet den Haupt-Transkriptionsstart, die dünnen Pfeile weisen auf drei Neben-Transkriptionsstartpunkte hin.

15

Fig. 6: cDNA Sequenz der humanen katalytischen Telomerase-Untereinheit (hTC; vgl. unsere anhängige Anmeldung PCT/EP/98/03468). Die dargestellte Sequenz entspricht SEQ ID NO 2.

20

Fig. 7: Strukturelle Organisation und Restriktionsmappe des humanen hTC-Gens und dessen 5'- und 3'-flankierende Region.

25

Exons sind als durchnummerierte schwarz ausgefüllte Rechtecke und Introns als nicht ausgefüllte Bereiche hervorgehoben. Nichttranslatierte Sequenzabschnitte in den Exons sind schraffiert. Die Translation startet in Exon 1 und endet in Exon 16. Restriktionsenzymchnittstellen sind wie folgt gekennzeichnet: S, SacI; X, XhoI. Die relative Anordnung der fünf Phagenklone (P2, P3, P5, P12, P17) und des Produktes aus dem „Genomic walking“ sind durch dünne Linien hervorgehoben. Wie durch

30

die Punkte gekennzeichnet, ist die Sequenz von Intron 16 nur teilweise entschlüsselt.

Fig. 8: HTL Splicevarianten.

5

10

15

A: Schematische Struktur der hTC mRNA Splicevarianten. Die vollständige hTC mRNA ist als grau unterlegtes Rechteck im oberen Bereich der Abb. dargestellt. Die 16 Exons sind entsprechend ihrer Größe dargestellt. Der Translationsstart (ATG) und das Stop-Codon, sowie das Telomerase-spezifische T-Motiv und die sieben RT-Motive sind hervorgehoben. Die hTC-Varianten sind in Deletions- und Insertionsvarianten unterteilt. In den Deletionen sind die fehlenden Exonsequenzen markiert. Die Insertionen sind durch zusätzliche weiße Rechtecke hervorgehoben. Größe und Herkunft der insertierten Sequenzen sind angegeben. Neu entstandene Stop-Codons sind markiert. Die Größe der Insertion von Variante INS2 ist unbekannt.

20

25

B: Exon Intron Übergänge der hTC-Splice-Varianten. Nichtgespligte 5'- und 3'-flankierte Sequenzen sind als weiße Rechtecke hervorgehoben. Die Herkunft der Exon und Intron Sequenzen ist angegeben. Intron und Exon Sequenzen sind in Kleinbuchstaben, bzw. in Großbuchstaben dargestellt. Die Donor und Akzeptor Sequenzen der Splice Stellen sind als graue Rechtecke unterlegt und deren Exon Intron Herkunft ist ebenfalls angegeben.

Fig. 9: Identifizierung des Transkriptionsstarts durch RT-PCR Analyse.

30

Die RT-PCR wurde mit cDNA-Bibliothek aus HL 60 Zellen und genomischer DNA als Positivkontrolle durchgeführt. Ein gemeinsamer 3'-Primer hybridisiert an eine Sequenzregion aus Exon 1. Die Position der verschiedenen 5' Primer in der kodierenden Region oder der 5'-flankierenden Region ist angegeben. In der Negativkontrolle wurde keine

Template-DNA in der PCR-Reaktion zugegeben. M: DNA-Größenmarker.

- Fig. 10: Nukleotidsequenz und Strukturmerkmale des hTC-Promotors.
- 5 11273 bp der 5'-flankierenden hTC Gensequenz, beginnend mit dem Translationsstartcodon ATG (+1) sind dargestellt. Die putative Region des Translationsstarts ist unterstrichen. Mögliche regulatorische Sequenzabschnitte innerhalb der 4000 bp stromaufwärts des Translationsstarts sind umrandet. Die dargestellte Sequenz entspricht SEQ ID NO 3.
- 10
- Fig. 11: Aktivität des hTC-Promotors in HEK-293 Zellen.
- Im oberen Bereich der Abbildung sind die ersten 5000 bp der 5'-flankierenden hTC Genregion schematisch dargestellt. Das ATG-Startcodon ist hervorgehoben. CpG reiche Inseln sind durch graue Rechtecke markiert. Auf der linken Abbildungsseite sind die Größen der hTC Promotor-Luziferase Konstrukte dargestellt. Das promotorlose pGL2-Basic Konstrukt und das SV40 Promotorkonstrukt pGL2-Pro wurden in jeder Transfektion als Kontrollen eingesetzt. Auf der rechten Abbildungsseite ist die relative Luziferaseaktivität der verschiedenen Promotorkonstrukte in HEK-Zellen als durchgehende Balken gezeigt. Die Standardabweichung ist angegeben. Die Zahlenwerte repräsentieren den Durchschnitt von zwei unabhängigen Experimenten, die in Duplikaten durchgeführt wurden.
- 15
- 20
- 25 Tab. 1: Exon Intron Übergänge des hTC-Gens
- Aufgelistet sind die Nukleotidsequenzen an den 3'- und 5' Spliceübergängen des hTC-Gens. Die Konsensussequenzen für Donor und Akzeptorsequenzen (AG und GT) sind durch graue Rechtecke unterlegt. Intronsequenzen (Kleinbuchstaben) und Exonsequenzen (Großbuchstaben), die die Spliceakzeptor- und Donorstellen flankieren sind
- 30
- gezeigt. Die Größe der Exons und Introns ist in bp angegeben.

Tab. 2: Potentielle Bindungsstellen für DNA-bindende Faktoren in der Nukleotidsequenz von Intron 2

5 Die Suche nach möglichen DNA-bindenden Faktoren (z.B. Transkriptionsfaktoren) wurde mit dem „Find Pattern“-Algorithmus aus dem „GCG Sequenz Analysis“ Programmpaket der „Genetics Computer Group“ (Madison, USA) durchgeführt. Aufgelistet sind die Abkürzungen der identifizierten DNA-bindenden Faktoren und deren Lokalisation in

10 Intron 2.

Tab. 1

3' Acceptor Sequence			5' Donor Sequence					
Intron	Exon	Exon No.	bp	Exon	Intron	Intron on	bp	
		No.				No.		
5' flankierende Region								
cagggcgttcccccgag	GTTCAGGCAGCGCTGCGT	1	281	CGCCCCCTCCTTCGCCAG	gtgggcctccccgggtcg	1	104	
catgtccttctctgttttaag	GTGTCTGCCTGAAGAGC	2	1354	TGGCTGCGCAGGAGCCCCAG	gtgaggaggtggtgcccgt	2	8616	
gaggggctctctatttgag	GGGTGGCTGTGTTCCGGC	3	196	TGCAAAAGCATTTGGAATCAG	gtactgtatccccacgcca	3	2089	
cccatgtgtcccccgag	ACAGCACTTGAAGAGGGTG	4	181	GTTCCGCAGAGAAAAAGAGG	gtggctgtgcttttggtta	4	687	
ctcgccctcactcacagag	GCCGAGCGTCTCACCTCGA	5	180	TGAGCTGTACTTTTGTCAAG	gtgggtgcgggggaccccc	5	494	
ccctctcctctgcgggag	GTGGATGTGACGGGGCGCGT	6	156	CAAGGCCTTCAAGAGCCAC	gtagggttcacgtgtgata	6	>4660	18
ctcccgctcgttttcgag	GTCTCTACCTTGACAGACC	7	96	TGCCGTCTCATCGAGCAG	gtcggggcactgccccgca	7	980	
ctgtgtcttccccgcccag	AGCTCTCCCTGAATGAGG	8	86	CCGTGCGCATCAGGGGCAA	gtgagtcaggtggcccaggt	8	2484	
gtattttcccttattttag	GTCTACGTCACAGTGCAG	9	114	CGGGGATTTCGGCGGGACGG	gtgaggcctctcttcccc	9	1984	
catggccccctctgcccag	GCTGCTCTGCGTTTGGTG	10	72	ACGGAAAAACCTTCTCTCAG	gtgaggccccgtgccgtgtg	10	1871	
attccccctgtgtctag	GACCTGTGTCCGAGGTGTC	11	189	TGCAGAGCGACTACTCCAG	gtgagcgcaactggcccga	11	3800	
tccttcttggcgactctag	CTATGCCCGACCTCCATC	12	127	CCTGTTTCTGGATTGCGAG	gtgagcaggctgatggtca	12	880	
ctgtccgcacatcctctag	GTGAACAGCGCTCCAGACGG	13	62	TCTGTCTGACGGGTACAG	gtgagccgccaccacaagggg	13	3187	
agcctctgttttccccag	GTTTCACGCATGTGTGCTG	14	125	CTGAAAGCGCAAGAACGCAG	gtatgtcaggtgcctggc	14	781	
ctgtgattttggcccccgag	GGATGTCGTGGGGGCCAA	15	138	CTGGGTCACTCAGACAG	gcaagtgtgggtggagccc	15	536	
	CCCAGACGCAGCTGAGTCG	16	664	TTTTTCAGTTTGTGAAAAA	3' flankierende Region			

Tab. 2

Faktoren	Lokalisation in Intron 2
C/EBP	2925
CRE.2	2749
Sp1	2378, 4094, 4526, 4787, 4835, 4995
AP-2 CS3	5099
AP-2 CS4	2213, 3699, 4667, 5878, 5938, 6059, 6180, 6496
AP-2 CS5	5350, 5798, 5880, 5940, 6061, 6182, 6375, 6498
PEA3	934, 2505
P53	2125
GR uteroglobin	848, 1487, 2956
PR uteroglobin	3331
Zeste-white	1577, 1619, 1703, 1745, 1787, 1829, 1871, 1913, 1955, 1997, 2039, 2081, 3518, 3709, 4765, 5014, 5055
GRE	846
MyoD-MCK right site/rev	447, 509, 558, 1370, 1595, 1900, 2028, 2099, 4557
MyoD-MCK left site	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
Ets-I CS	6408
API	3784, 4406
CREB	2801
GATA-1	839, 1390, 3154
c-Myc	108, 118, 453, 1566, 1608, 1692, 1734, 1818, 1902, 1986, 2372, 2460, 2720, 3491, 5030
CACCC site	991
CCAAT site	1224
CCAC box	992
CAAT site	463, 2395
Rb site	992, 4663
TATA	3650
CDEI	106, 1564, 1606, 1690, 1732, 1816, 1900, 1984

### Beispiele

Das menschliche Gen für die katalytische Telomerase Untereinheit (ghTC), sowie die 5' und 3' liegenden Bereiche dieses Gens wurden kloniert, der Startpunkt der Transkription bestimmt, potentielle Bindungsstellen für DNA-bindende Proteine identifiziert, sowie aktive Promotorfragmente aufgezeigt. Die Sequenz der hTC-cDNA (Fig. 6) ist bereits in unserer ebenfalls anhängigen Anmeldung PCT/EP/98/03468 beschrieben. Wenn nicht gesondert erwähnt, beziehen sich sämtliche Angaben zur cDNA-Position auf diese Sequenz.

#### Beispiel 1

Durch eine genomische Southern Blot-Analyse wurde bestimmt, ob ghTC im menschlichen Genom ein Einzelgen darstellt oder mehrere Loci für das hTC-Gen bzw. eventuell auch ghTC-Pseudogene existieren.

Hierzu wurde ein kommerziell erhältlicher Zoo-Blot der Firma Clontech einer Southern Blot-Analyse unterzogen. Dieser Blot enthält 4 µg Eco RI geschnittene genomische DNA von neun verschiedenen Spezies (Mensch, Affe, Ratte, Maus, Hund, Rind, Kaninchen, Huhn und Hefe). Mit Ausnahme von Hefe, Huhn und Mensch wurde die DNA aus Nierengewebe isoliert. Die humane genomische DNA wurde aus Plazenta isoliert und die genomische DNA aus Huhn wurde aus Lebergewebe aufgereinigt. Im Autoradiogramm in Fig. 1 wurde als radioaktiv-markierte Sonde ein etwa 720 bp langes hTC-cDNA Fragment, isoliert aus der hTC cDNA, Variante Del2 (Position 1685 bis 2349 plus 2531 bis 2590 der Fig. 6 [Deletion 2; vergl. Beispiel 5 der Fig. 8]), eingesetzt. Die experimentellen Bedingungen für die Hybridisierung und die Waschschrte des Blots erfolgten in Anlehnung an Ausubel *et al.* (1987).

Im Fall der humanen DNA erkennt die Sonde zwei spezifische DNA-Fragmente. Das kleinere, etwa 1,5 bis 1,8 kb lange Eco RI-Fragment geht wahrscheinlich auf zwei

Eco RI-Schnittstellen in einem Intron der ghTC-DNA zurück. Aufgrund dieses Ergebnisses ist davon auszugehen, daß nur ein singuläres ghTC-Gen im menschlichen Genom vorliegt.

## 5 Beispiel 2

Zur Isolierung der 5' flankierenden hTC-Gensequenz wurden ca  $1,5 \times 10^6$  Phagen einer humanen genomischen Plazenta-Genbibliothek (EMBL 3 SP6/T7 der Firma Clontech, Bestellnummer HL1067j) auf Nitrozellulosefilter (0,45 µm; Fa. Schleicher und Schuell) nach Angaben des Herstellers mit einem radioaktiv markierten, etwa  
10 500 bp langen 5'-hTC-cDNA Fragment (Position 839 bis 1345 der Fig. 6) hybridisiert. Die Nitrozellulosefilter wurden zunächst in 2 x SSC (0,3 M NaCl; 0,5 M Tris-HCl, pH 8,0) und anschließend in einer Prähybridisierungslösung (50 % Formamid; 5 x SSPE, pH 7,4; 5 x Denhards-Lösung; 0,25 % SDS; 100 µg/ml  
15 Heringssperma-DNA) zwei Stunden bei 42°C inkubiert. Für die Hybridisierung über Nacht wurde die Prähybridisierungslösung mit  $1,5 \times 10^6$  cpm/ml Lösung denaturierter, radioaktiv markierter Probe ergänzt. Unspezifisch gebundene, radioaktive DNA wurde unter stringenten Bedingungen, d.h. durch drei fünfminütige Waschschritte mit 2 x SSC; 0,1 % SDS bei 55 bis 65 °C entfernt. Die Auswertung erfolgte  
20 durch Autoradiographie der Filter.

Die in dieser Primäruntersuchung identifizierten Phagenklone wurden aufgereinigt Ausubel *et al.* (1987). In weitergehenden Analysen stellte sich ein Phagenklon P12 als potentiell positiv heraus. Eine λ-DNA Präparation dieses Phagens Ausubel *et al.*  
25 (1987) und der nachfolgende Restriktionsverdau mit Enzymen, die das genomische Insert in Fragmenten freisetzen, zeigte, daß dieser Phagenklon ein ca. 15 kb Insert im Vektor enthält (Fig. 2).

Zur Isolierung der vollständigen hTC-Gensequenz wurden in unabhängigen  
30 Experimenten jeweils 1 bis  $1,5 \times 10^6$  Phagen mit jeweils verschiedenen radioaktiv markierten Sonden wie oben beschrieben durchmustert.

Die in diesen Primäruntersuchungen identifizierten, für die entsprechenden Sonden positiven Phagenklone wurden aufgereinigt. Der Phagenklon P17 wurde mit einem etwa 250 bp langen hTC-cDNA Fragment (Position 1787 bis 2040 der Fig. 6) gefunden. Der Phagenklon P2 wurde mit einem etwa 740 bp langen hTC-cDNA Fragment (Position 1685 bis 2349 plus 2531 bis 2607 der Fig. 6 [Deletion 2; vergl. Beispiel 5]) identifiziert. Die Phagenklone P3 und P5 wurden mit einem 420 bp langen 3' hTC-cDNA Fragment (Position 3047 bis 3470 der Fig. 6) gefunden. Nach  $\lambda$ -DNA Präparation dieser Phagen und nachfolgendem Restriktionsverdau mit Enzymen, die das genomische Insert in Fragmenten freisetzen, wurden die Inserts in Plasmide umklont (Beispiel 4).

### **Beispiel 3**

Um zu untersuchen, ob auch das 5'-Ende der hTC-cDNA im Insert des rekombinanten Phagenklons P12 vorliegt, wurde  $\lambda$ -DNA dieses Klons in einer Southern Blot Analyse mit einem radioaktiv markierten etwa 440 bp langen hTC-cDNA Fragment (Position 1 bis 440 der Fig. 6) aus dem extremen 5'-Bereich hybridisiert (Fig. 3).

Da die isolierte  $\lambda$ -DNA des positiven Klons auch mit dem extremen 5'-Ende der hTC-cDNA hybridisiert, enthält dieser Phage wahrscheinlich auch den das ATG-Startcodon flankierenden 5'-Sequenzbereich.

### **Beispiel 4**

Um das gesamte 15 kb lange Insert des positiven Phagenklons P12 in Teilfragmenten umzuklonieren und anschließend zu sequenzieren, wurden zum DNA-Verdau Restriktionsendonukleasen ausgewählt, die zum einen das gesamte Insert aus EMBL3 Sp6/T7 freisetzen (vgl. Beispiel 2) und zusätzlich im Insert schneiden.

Insgesamt wurden ein etwa 8,3 und ein etwa 6,5 kb langes Xho I-Subfragment sowie ein etwa 8,5, ein etwa 3,5 und ein etwa 3 kb langes Sac I-Teilfragment in den Vektor pBluescript KS(+) (Fa. Stratagene) umkloniert. Durch Sequenzanalyse dieser Fragmente wurde die Nukleotidsequenz von 5123 bp 5'-flankierenden des ghTC-Genbereichs, ausgehend vom ATG-Startcodon bestimmt (Fig. 4; entsprechend SEQ ID NO 1). In der Fig. 4 sind die ersten (ausgehend vom ATG-Startcodon) 5123 bp dargestellt. In der Fig. 10 (entsprechend SEQ ID NO 3) die gesamte klonierte 5' Sequenz.

Um das gesamte ca. 14,6 kb große Insert des Phagenklons P17 in Teilfragmenten umzuklonieren, wurden zum DNA-Verdau Restriktionsendonukleasen ausgewählt, die zum einen das gesamte Insert aus EMLB3 Sp6/T7 freisetzen und zusätzlich einige Male im Insert schneiden. Durch Kombinationsverdau mit den Enzymen XhoI und BamHI wurden ein 7,1 kb, ein 4,2 kb und ein 1,5 kb großes XhoI-BamHI-Fragment sowie ein 1,8 kb großes BamHI-Fragment subkloniert. Der Kombinations-Restriktionsverdau mit den Enzymen XhoI und XbaI führte zur Klonierung von einem 6,5 kb großen XhoI-XbaI-Fragment, einem 6,5 kb und einem 1,5 kb großem XhoI-Fragment.

Die Umklonierung des ca. 17,9 kb großem Inserts des Phagenklons P2 in Subfragmente erfolgte durch Verdau mit dem Restriktionsenzym XhoI. Insgesamt wurde ein 7,5 kb, ein 6,4 kb sowie ein 1,6 kb langes XhoI-Subfragment kloniert. Durch Verdau mit dem Restriktionsenzym SacI wurde zusätzlich ein 4,8 kb, ein 3 kb, ein 2 kb sowie ein 1,8 kb großes SacI-Fragment subkloniert.

Das ca. 13,5 kb große Insert des Phagenklons P3 wurde durch Verdau mit den Restriktionsenzymen SacI bzw. XhoI subkloniert. Dabei wurden ein 3,2 kb, ein 2 kb, ein 0,9 kb, ein 0,8 kb, ein 0,65 kb und ein 0,5 kb langes SacI-Subfragment sowie ein 6,5 kb und ein 4,3 kb langes XhoI-Subfragment erhalten.

30

Die Subklonierung des ca. 13,2 kb großen Inserts des Phagenklons P5 erfolgte durch Verdau mit den Restriktionsenzymen SacI bzw. XhoI. Insgesamt wurden SacI-Fragmente von 6,5 kb, 3,3 kb, 3,2 kb, 0,8 kb und 0,3 kb Größe sowie XhoI-Fragmente von 7 kb und 3,2 kb Größe subkloniert.

5

Zur Klonierung des 3' von Phagenklon P17 und 5' von Phagenklon P2 gelegenen hTC-genomischen Sequenzbereichs wurden 3 Genomic Walkings mit Hilfe des GenomeWalker™ Kits der Firma Clontech (Katalognummer K1803-1) und verschiedenen Primerkombinationen durchgeführt. In einem Endvolumen von 50 µl wurde 1 µl humaner GenomeWalker Library HDL (Fa. Clontech) mit 10 pmol dNTP-Mix versetzt und in 1xKlen Taq PCR-Reaktionspuffer und 1xAdvantage Klen Taq Polymerase Mix (Fa. Clontech) eine PCR-Reaktion durchgeführt. Als Primer wurden 10 pmol eines internen genspezifischen Primers sowie 10 pmol des Adaptor Primers AP1 (5'-GTAATACGACTCACTATAGGGC-3'; Fa. Clontech) zugefügt.

15 Die PCR wurde als Touchdown-PCR in 3 Schritten durchgeführt. Zunächst wurde über 7 Zyklen für 20 sec bei 94°C denaturiert und anschließend für 4 min bei 72°C die Primer angelagert und die DNA-Kette verlängert. Es folgten 37 Zyklen bei denen für 20 sec die DNA bei 94°C denaturiert wurde, die anschließende Primerverlängerung aber für 4 min bei 67°C erfolgte. Abschließend folgte eine Kettenverlängerung für 4 min bei 67°C. Im Anschluß an diese erste PCR wurde das PCR-Produkt 1:50 verdünnt. Ein µl dieser Verdünnung wurde in einer zweiten „nested“ PCR zusammen mit 10 pmol dNTP-Mix in 1xKlen Taq PCR-Reaktionspuffer und 1xAdvantage Klen Taq Polymerase-Mix sowie 10 pmol eines „nested“ genspezifischen Primers und 10 pmol des „nested“ Marathon Adaptor Primers AP2 (5'-

25 ACTATAGGGCACGCGTGGT-3'; Fa. Clontech) eingesetzt. Die PCR-Bedingungen entsprachen den in der ersten PCR gewählten Parametern. Als einzige Ausnahme wurden im ersten PCR-Schritt statt 7 Zyklen nur 5 Zyklen gewählt und im zweiten PCR-Schritt statt 37 Zyklen nur 24 Zyklen durchlaufen. Produkte dieser Nested-GenomicWalking-PCR wurden in den TA-Cloning Vektor pCRII der Fa. InVitrogen

30 kloniert.

Im ersten Genomic Walking wurde der genspezifische Primer C3K2-GSP1 (5'-GACGTGGCTCTTGAAGGCCTTG-3') sowie der „nested“ genspezifische Primer C3K2-GSP2 (5'-GCCTTCTGGACCACGGCATACC-3') zusammen mit der HDL-Library 4 eingesetzt und ein 1639 bp langes PCR-Fragment erhalten. Im zweiten  
5 Genomic Walking wurde mit dem genspezifischen Primer C3F2 (5'-CGTAGTTGAGCACGCTGAACAGTG-3') und dem „nested“ genspezifischen Primer C3F (5'-CCTTCACCCTCGAGGTGAGACGCT-3') aus der HDL-Library 4 ein PCR-Fragment von 685 bp Länge amplifiziert. Der dritte Genomic Walking  
10 Ansatz führte unter Einsatz des genspezifischen Primers DEL5-GSP1 (5'-GGTGGATGTGACGGGCGCGTACG-3') und des „nested“ genspezifischen Primers C5K-GSP1 (5'-GGTATGCCGTGGTCCAGAAGGC-3') zur Klonierung eines 924 bp PCR-Fragments aus der HDL-Library 1. Insgesamt wurden durch dieses Genomic Walking-Verfahren 2100 bp der 3' von Phagenklon P17 gelegenen genomischen hTC-Region identifiziert (s. Fig. 7).

15 Die subklonierten Fragmente sowie die Genomic Walking-Produkte wurden einzelsträngig sequenziert. Unter Verwendung der Lasergene Biocomputing Software (DNASTAR Inc. Madison, Wisconsin, USA) wurden überlappende Bereiche identifiziert und Contigs gebildet. Insgesamt wurden aus den gesammelten  
20 Sequenzen der Phagenklone P12, P17, P2, P3 und P5 sowie den Sequenzdaten aus dem Genomic Walking 2 große Contigs zusammengestellt. Contig 1 besteht aus Sequenzdaten von Phagenklon P12, P17 und den Sequenzdaten aus dem Genomic Walking. Contig 2 wurde aus den Sequenzen von Phagenklon P2, P3 und P5 zusammengesetzt. Überlappende Phagenklonbereiche sind in Fig. 7 schematisch  
25 dargestellt. Die Sequenzdaten der 2 Contigs sind nachfolgend dargestellt. Das ATG Startcodon in Contig 1 ist unterstrichen. Das TGA Stopcodon ist in Contig 2 unterstrichen.

## Contig1:

5	ACTTGAGCCC	AAGAGTTC	GGCTACGG	AGCCATGATT	GCAACACCAC	ACGCCAGCCT	TGGTGACAGA	70
	ATGAGACCCT	GTCTCAAAA	AAAAA	AATTGAAATA	ATATAAGCA	TCTTCTCTGG	CCACAGTGGA	140
	ACAAACCAG	AAATCAACAA	CAAGAGGAAT	TTTGAAACT	ATACAAACAC	ATGAAATTA	AACATATATAC	210
	TTCTGAATGA	CCAGTGAGTC	AATGAAGAA	TTAAAAAGGA	AATTGAAAAA	TTTATTTAAG	CAATATGATA	280
	CGGAAACATA	ACCTCTCAAA	ACCCACGGTA	TACAGCAAAA	GCAGTGCTAA	GAAGGAAGTT	TATAGCTATA	350
	AGCAGCTACA	TCAAAAAAGT	AGAAAAGCCA	GGCCGAGTGG	CTCATGCCCTG	TAATCCCAGC	ACTTTGGGAG	420
10	GCCAGGGGG	GCAGATCGCC	TGAGGTCAG	AGTTCGAGAC	CAGCCTGACC	AACACAGAGA	AACCTTGTGC	490
	CTACTAAAA	TACAAATTA	GCTGGGCATG	GTGGCAGATG	CCTGTAATCC	CAGCTACTCG	GGAGGCTGAG	560
	GCAGGATAAC	CGCTTGAAAC	CAGGAGGTGG	AGGTTGCGGT	GAGCCGGGAT	TGGCCCATTG	GACTCCAGCC	630
	TGGGTAAACA	GAGTGAAACC	CTGTCTCAAG	AAAAA	AAGTAGAAAA	ACTTAAAAAT	ACAACCTAAT	700
	GATGCACCTT	AAAGAACTAG	AAAAGCAAGA	GCAAACTAAA	CCTAAATTTG	GTAAGAAAA	AGAAATAATA	770
15	AAGATCAGAG	CAGAAATAAA	TGAACTGAA	AGATAACAAT	ACAAAAGATC	AACAAATTA	AAAGTTGGTT	840
	TTTTGAAAG	ATAAACAAA	TTGACAAACC	TTTGCCAG	CTAAGAAAA	AGGAAAGAAG	ACCTAAATAA	910
	ATAAAGTCAG	AGATGAAAA	AGAGACATTA	CAACTGTATC	CACAGAAATT	CAAGGATCA	CTGAGGCTA	980
	CTGTAGCAA	CTGTACACTA	ATAAATTGAA	AAACCTAGAA	AAAAATAGATA	AATTCTTAGA	TGCATCAAC	1050
	CTACCAAGAT	TGAACCATGA	AGAAATCCAA	AGCCCAACCA	GACCAATTAAC	AATAATGGGA	TAAAGGCAT	1120
20	AATAAAAGT	CTCCTAGCAA	AGAGAAGCCC	AGGACCCAAT	GGCTTCCCTG	CTGGATTTTA	CCAATCATT	1190
	AAAGAAGAT	GAATTCCAAT	CCTACTCAAA	CTATTCTGAA	AAATAGAGGA	AAGAATACTT	CCAACTCAT	1260
	TCTACATGGC	CAGTATTACC	CTGATTCCAA	AACCAGACAA	AAACACATCA	AAAACAAACA	AACAAAAA	1330
	CAGAAAGAAA	GAAAACACAA	GGCCAATATC	CCTGATGAAT	ACTGATACAA	AAATCCTCAA	CAAAACACTA	1400
	GCAAAACAAA	TTAAACAACA	CCTTCGAAAG	ATCATTCTAT	GTGATCAAGT	GGGATTATT	CCAGGGATGG	1470
25	AAGGATGGTT	CAACATATGC	AAATCAATCA	ATGTGATACA	TCATCCCAAC	AAAATGAAGT	ACAAAACATA	1540
	TATGATTATT	TCACTTTATG	CAGAAAAAGC	ATTTGATAAA	ATTCTGCACC	CTTCATGATA	AAAACCCTCA	1610
	AAAAACGAG	TATACAAGAA	ACATACAGCC	CAGGCACAGT	GGCTCACACC	TGGGATCCCA	GCACTCTGGG	1680
	AGGCCAAGGT	GGGATGATTG	CTTGGGCCCA	GGAGTTTGAG	ACTAGCCTGG	GCAACAAAT	GAGACCTGGT	1750
	CTACAAAAA	CTTTTTTAA	AAATTAGCCA	GGCATGATGG	CATATGCCCTG	TAGTCCACAG	TAGTCTGGAG	1820
30	GCTGAGGTGG	GAGAACTACT	TAAGCCTAGG	AGGTCGAGGC	TGCAGTGAGC	CATGAACATG	TCACCTGACT	1890
	CCAGCCTAGA	CAACAGAAC	AGACCCCACT	GAATAAGAAG	AAGGAGAAGG	AGAAGGGAGA	AGGGAGGGAG	1960
	AAGGGAGGAG	GAGGAGAAGG	AGGAGGTGGA	GGAGAAGTGG	AAGGGGAAGG	GGAAAGGAAA	GAGGAAGAAG	2030
	AAGAAACATA	TTTCAACATA	ATAAAGCCC	TATATGACAG	ACCGAGGTAG	TATTATGAGG	AAAACTGAA	2100
	AGCCTTTCT	CTAAGATCTG	GAAATGACA	AGGGCCCACT	TTCACCACCTG	TGATTCAACA	TAGTACTAGA	2170
35	AGTCTAGCT	AGAGCAATCA	GATAAGAGAA	AGAAATAAAA	GGCATCCAAA	CTGGAAAGGA	AGAACTCAAA	2240
	TTATCCTGTT	TGCAGATGAT	ATGATCTTAT	ATCTGGAAAA	GACTTAAGAC	ACCACTAAAA	AACATATTAGA	2310
	GCTGCAATTT	GGTACAGCAG	GATACAAAT	CAATGTACAA	AAATCAGTAG	TATTCTTATA	TTCCAACAGC	2380
	AAACAATCTG	AAAAAGAAC	CAAAAAGCA	GCTACAAATA	AAATTAACA	GCTAGGAATT	AACCAAGAA	2450
	GTGAAGATC	TCTACATGA	AACTATAAA	ATGTTGATA	AAGAAATTGA	AGAGGGCACA	AAAAAGAAA	2520
40	AGATATTCCA	TGTTCATAGA	TTGGAAGAA	AAATCTGTT	AAATGTCCA	TACTACCCAA	AGCAATTTAC	2590
	AAATTCATG	CAATCCCTAT	TAAATACTA	ATGACGTTCT	TCACAGAAAT	AGAAGAAACA	ATTCTAAGAT	2660
	TTGTACAGAA	CCACAAAAGA	CCCAGAAATG	CCAAAGCTAT	CCTGACCAAA	AAGAACAAAA	CTGGAAGCAT	2730
	CACATTACCT	GACTTCAAT	TATACTACAA	AGCTATAGTA	ACCAAACTA	CATGGTACTG	GCATAAAAAC	2800
	AGATGAGACA	TGGACCAGAG	AGACAGAATA	GAGAATCCAG	AAACAAATCC	ATGCATCTAC	AGTGAACCTA	2870
45	TTTTTGACAA	AGGTGCCAAG	AACATACTTT	GGGAAAAGAA	TAATCTCTTC	AATAAATGGT	GCTGGAGGAA	2940
	CTGGATATCC	ATATGCAAAA	TAACAATACT	AGAACTCTGT	CTCTCACCAT	ATACAAAAGC	AAATCAAAAT	3010
	GGATGAAAGG	CTTAAATCTA	AAACCTCAAA	CTTTGCAACT	ACTAAAAGAA	AACACCGGAG	AACCTCTCCA	3080
	GGACATTTGA	GTGGGCAAG	ACTTCTTAG	TAATTCCTGT	CAGGCACAGG	CAACCAAGC	AAAAACAGAC	3150
	AAATGGGATC	ATATCAAGTT	AAAAAGCTTC	TGCCACGCAA	AGGAAACAA	CAACAAAGAG	AAGAGACAA	3220
50	CCACAGAAATG	GGAGAATATA	TTTGCAAACT	ATTCATCTAA	CAAGGAATTA	ATAACCAGTA	TATATAAGGA	3290
	GCTCAAACTA	CTCTATAAGA	AAAACACCTA	ATAAGCTGAT	TTTCAAAAT	AAGCAAAAGA	TCTGGGTAGA	3360
	CATTTCTCAA	AATAAGTCAT	ACAAATGGCA	AACAGGCATC	TGAAATGTG	CTCAACACCA	CTGATCATCA	3430
	GAGAAATGCA	AATCAAACT	ACTATGAGAG	ATCATCTCAT	CCCAGTTAAA	ATGGCTTTTA	TTCAAAAGAC	3500
	AGGCAATAAC	AAATGCCAGT	GAGGATGTGG	ATAAAGGAA	ACCCTTGGAC	ACTGTTGGTG	GGAATGGAAA	3570
55	TTGCTACCAC	TATGGAGAAC	AGTTTGAAAG	TTCTCAAAA	AACTAAAAAT	AAAGCTACCA	TACAGCAATC	3640
	CCATTGTCTAG	GTATATACTC	CAAAAAAGGG	AATCAGTGTA	TCAACAAGCT	ATCTCCACTC	CCACATTAC	3710
	TGCAGCACTG	TTCATAGCAG	CCAAGGTTTG	GAAGCAACCT	CAGTGTCCAT	CAACAGACGA	ATGGAAAAAG	3780
	AAAATGTGGT	GCACATACAC	AATGGAGTAC	TACGCAGCCA	TAAAAAGAA	TGAGATCCTG	TCAGTTGCAA	3850
	CAGCATGGGG	GGCACTGGTC	AGTATGTTAA	GTGAAATAAG	CCAGGCACAG	AAAGACAAAC	TTTTCATGTT	3920
60	CTCCCTTACT	TGTGGGAGCA	AAAATTAATA	CAATTGACAT	AGAAATAGAG	GAGAATGGTG	GTTCTAGAGG	3990
	GGTGGGGGAC	AGGGTGACTA	GAGTCAACAA	TAATTTATTG	TATGTTTTAA	AATAACTAAA	AGAGTATAAT	4060
	TGGGTTGTTT	GTAACACAAA	GAAAGGATAA	ATGCTTGAAG	GTGACAGATA	CCCATTAC	CCTGATGTGA	4130
	TTATTACACA	TTGTATGCCCT	GTATCAAAAT	ATCTCATGTA	TGCTATAGAT	ATAAACCTTA	CTATATTAAA	4200
	AATTAATAAT	TTAATGGCCA	GGCACGGTGG	CTCATGTCCG	TAATCCCAGC	ACTTTGGGAG	GCCGAGGCGG	4270
65	GTGGATCACC	TGAGGTCCAG	AGTTTGAAC	CAGTCTGGCC	ACCATGATGA	AACCTGTCT	CTACTAAGA	4340
	TACAAAAAT	AGCCAGGCGT	GGTGGCACAT	ACCTGTAGTC	CCAATCTACT	AGGAGGCTGA	GACAGGAGAA	4410
	TTGCTTGAAC	CTGGAGGCGC	GAGGTTGCAG	TGAGCCGAGA	TCATGCCACT	GCACTGCAGC	CTGGGTGACA	4480
	GAGCAAGACT	CCATCTCAAA	ACAAAAACAA	AAAAAGAAAG	ATTAAAAATTG	TAATTTTTAT	GTACCGTATA	4550
70	AATATATACT	CTACTATATT	AGAAGTTAAA	AATTAAAAACA	ATTATAAAG	GTAATTAAAC	ACTTAATCTA	4620
	AAATAGAAC	AATGTATGTG	GGGTTTCTAG	CTTCTGAAGA	AGTAAAAAGT	ATGGCCACGA	TGGCAGAAAT	4690
	GTGAGGAGGG	AACAGTGGAA	GTTACTGTGG	TTAGACGCTC	ATACTCTCTG	TAAGTGACTT	AATTTAAC	4760
	AAAGACAGGC	TGGGAGAAGT	TAAAGAGGCA	TTCTATAAGC	CCTAAAAACA	CTGCTAATAA	TGGTGAAAGG	4830
	TAATCTCTAT	TAATTACCAA	TAATTACAGA	TATCTCTAAA	ATCGAGCTGC	AGAATTGGCA	CGCTGTGATCA	4900
	CACCGTCTCT	TCATTACAGG	TGCTTTTTT	CTTGTGTGCT	TGGAGATTTT	CGATTGTGTG	TCTGTGTTT	4970
75	GTTAAACTTA	ATCTGTATGA	ATCCTGAAAC	GAAAAATGGT	GGTGATTTC	TCCAGAGAA	TTAGAGTACC	5040
	TGCCAGGAAG	CAGGTGGCTC	TGTGGACCTG	AGCCACTTCA	ATCTTCAAGG	GTCTCTGGCC	AAGACCCAGG	5110

	TGCAAGGCAG	AGGCCTGATG	ACCCGAGGAC	AGGAAAGCTC	GGATGGGAAG	GGGCGATGAG	AAGCCTGCCT	5180
	CGTTGGTGAG	CAGCGCATGA	AGTGCCCTTA	TTTACGCTTT	GCAAAGATTG	CTCTGGATAC	CATCTGGAAA	5250
	AGGCGGCCAG	CGGGAATGCA	AGGAGTCAGA	AGCCTCCTGC	TCAAACCCAG	GCCAGCAGCT	ATGGCGCCCA	5320
	CCCGGGCGTG	TGCCAGAGGG	AGAGGAGTCA	AGGCACCTCG	AAGTATGGCT	TAAATCTTTT	TTTCACCTGA	5390
5	AGCAGTGACC	AAGGTGTATT	CTGAGGGAAG	CTTGAGTTAG	GTGCTTCTTT	TAAACAGAA	AGTCATGGAA	5460
	GCACCTTCT	CAAGGGAAAA	CCAGACGCCC	GCTCTGCGGT	CATTACCTC	TTTCTCTCT	CCCTCTCTTG	5530
	CCCTCGCGGT	TTCTGATCGG	GACAGAGTGA	CCCCCGTGGA	GCTTCTCCGA	GGCCGTGCTG	AGGACCTCT	5600
	TGCAAGGGG	TCCACAGACC	CCCGCCTGG	AGAGAGGAGT	CTGAGCCTGG	CTTAATAACA	AACCTGGATG	5670
	TGGCTGGGG	CGGACAGCGA	CGCGGGGATT	CAAGAGCTTA	ATTCCATGAG	TAAATTC AAC	CTTTCCACAT	5740
10	CCGAATGGAT	TTGGATTTTA	TCTTAATATT	TTCTTAAATT	TCATCAAATA	ACATTACAGGA	CTGCAGAAAT	5810
	CCAAAGGCGT	AAAACAGGAA	CTGAGCTATG	TTTGCCAAAG	TCCAAGGACT	TAATAACCAT	GTTCAGAGGG	5880
	ATTTTTCGCC	CTAAGTACTT	TTTATTGGTT	TTTATAAGGT	GGCTTAGGGT	GCAAGGGAAA	GTACACGAGG	5950
	AGAGGCTG	GCGGCAGGGC	TATGAGCACG	GCAGGGCCAC	CGGGAGAGAG	GTCCCCGGCC	TGGGAGGCTG	6020
	ACAGCAGGAC	CACTGACCGT	CCTCCTCTGG	AGCTGCCACA	TTGGGCAACG	GGAAGCGGGC	CACGCTCGCT	6090
15	GTGACTCAGG	ACCCCATACC	GGCTTCCTGG	GGCCACCCAC	ACTAACCCAG	GAAGTCACGG	AGCTCTGAAC	6160
	CCGTGGAAAC	GAACATGACC	CTTGCTCTGC	TGCTTCCCTG	GGTGGGTCAA	GGGTAAATGAA	GTGCTGTGCA	6230
	GGAAATGGCC	ATGTAAATTA	CACGACTCTG	CTGATGGGGA	CCGTTCCTTC	CATCATTATT	CATCTTCACC	6300
	CCCAAGGACT	GAATGATTCC	AGCAACTTCT	TCGGGTGTGA	CAAGCCATGA	CAAACTCAG	TACAAACACC	6370
20	ACTCTTTTAC	TAGGCCACAC	GAGCAGCGSC	CACACCCCTG	ATATATTAAG	AGTCAGGAG	AGATGAGGCT	6440
	GCTTTACG	ACCAGGCTGG	GGTGACAACA	CCGCTGAAC	AGTCTGTTC	TCTAGACTAG	TAGACCTTGG	6510
	CAGGCATCC	CCCAGATTCT	AGGGCTCGT	TGCTGCTTCC	CGAGGGCGCC	ATCTGCCCTG	GAGACTCAGC	6580
	CTGGGTGCC	ACACTGAGGC	CAGCCCTGTC	TCCACACCTT	CCGCTCCAG	GCCTCAGCTT	CTCCAGCAGC	6650
	TTCTTAAAC	CTGGGTGGGG	CGTGTTCAG	CGCTACTGTC	TCACCTGTCC	CACCTGTGCT	TGCTCAGGCT	6720
	ACGTAGCTCG	CACGGTTCCT	CCTCAGCTGG	GGTGTCTGTC	TCCTTCCCCA	ACACTCAGCT	CGGTGAAGG	6790
25	GAGGAGATT	TGCGCTTCCC	AGACTGGCT	CTCTGAGCCT	GAACCTGGCT	CGTGGCCCCC	GTGCGAGGTT	6860
	CCTGGCGTCC	GGCTGCACGC	TGACCTCCAT	TTCCAGGCGC	TCCCGCTCTC	CTGTCTCTG	CCGGGGCCTG	6930
	CCGGTGTGTT	CTTCTGTTTC	TGTGCTCTT	TCCACGTCCA	GCTGCGTGTG	TCTCTGCCCC	CTAGGGTCTC	7000
	GGGGTTTTTA	TAGGCATAGG	ACGGGGGCGT	GGTGGGCCAG	GGCGCTCTTG	GGAAATGCAA	CATTTGGGTT	7070
30	TGAAAGTAGG	AGTGCCTGTC	CTCAGCTAGG	TCCACGGGCA	CAGGCCTGGG	GATGGAGCCC	CCGCCAGGGA	7140
	CCGCGCCTTC	TCTGCCACGC	ACTTTCCTCG	CCCCCTCCCT	CTGGAACACA	GAGTGGCAGT	TTCCACAAGC	7210
	ACTAAGCATC	CTCTTCCCAA	AAGACCCAGC	ATTGGCACCC	CTGGACATTT	GCCCCACAGC	CCTGGGAATT	7280
	CACGTGACTA	CGACATCAT	GTACACACTC	CCGTCCACGA	CCGACCCCGG	CTGTTTTATT	TTAATAGCTA	7350
	CAAAGCAGGG	AAATCCCTGC	TAAATGTCTC	TTTAAACAAAC	TGGTTAAACA	AACGGGTCCA	TCCGCACGGT	7420
35	GGACAGTTCC	TCACAGTGAA	GAGGAACATG	CCGTTTATAA	AGCCTGCAGG	CATCTCAAGG	GAATTACGGT	7490
	GAGTCAAAA	TGCCACCTCC	ATGGGATACG	TACGCAACAT	GCTCAAAAAG	AAAGAAATTC	ACCCCATGGC	7560
	AGGGGAGTGG	TAGGGGGGTT	TAGGACGCTG	GGGGGGGCGA	GCTGGGGGCT	ACTGCACGCA	CCTTTTACTA	7630
	AAGCCAGTTT	CCTGGTCTG	ATGGTATTGG	CTCAGTTATG	GGAGACTAAC	CATAGGGGAG	TGGGGATGGG	7700
	GGAACCCGGA	GGCTGTGCCA	TCTTTGCCAT	GCCCGAGTGT	CCTGGGCAGG	ATAATGCTCT	AGAGATGCCC	7770
40	ACGTCCTGAT	TCCCCCAAAC	CTGTGGACAG	AACCCGCCCC	GCCCCAGGGC	CTTTGCAGGT	GTATCTCCG	7840
	TGAGGACCTT	GAGGTCTGGG	ATCCTTCGGG	ACTACCTGCA	GGCCCCGAAA	GTAATCCAGG	GGTTCCTGGG	7910
	AGAGGCGGGC	AGGAGGGTCA	GAGGGGGGCA	GCCTCAGGAC	GATGGAGGCA	GTCAGTCTGA	GGCTCAAAA	7980
	GGAGGGAGGG	CCTCGAGCCC	AGGCTGCAA	GCCTCCAG	AAGCTGGAAA	AAGCGGGGAA	GGGACCTCC	8050
	ACGGAGCCTG	CAGCAGGAAG	GCAAGGCTGG	CCCTTAGCCC	ACCAGGGCCC	ATCGTGACC	TCCGGCTCC	8120
45	TGCCCATAGG	AGGCACTCG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	ATTTGCAGAA	GCAACAGGAA	8190
	AGCCCTGCAC	TGTGAATCTA	GGATTATTTC	AAAACAAAAG	TTTACAGAAA	CATCCAAGGA	CAGGGCTGAA	8260
	GTGCCCTCCG	GCAAGGGCAG	GGCAGGCACG	AGTGATTTTA	TTTAGCTATT	TTATTTTATT	TACTTACTTT	8330
	CTGAGACAGA	GTTATGCTCT	TGTTGCCAG	AGCGGATGCA	TCTTGGCTCA	TCTTGGCTCA	CTGCAACCTC	8400
	CGTCTCCTGG	GTTCAAGCAA	TTCTCGTGCC	TCAGCCTCCC	AAGTAGCTGG	GATTTACAGC	GTGCACACC	8470
50	ACACCCGGCT	AATTTGTAT	TTTTAGTAGA	GATGGGCTTT	CACCATGTTG	GTCAGCTGA	TCTCAAAATC	8540
	CTGACCTCAG	GTGATCCGCC	CACCTCAGCC	TCCCAAGATG	CTGGGATTAC	AGGCATGAGC	CAGCTGACCT	8610
	GGCCTATTTA	AGCATTTTAA	AATCTCCCTG	GGCTCAAGTC	ACACCCACTG	GTAAGGAGTT	CATGGAGTTC	8680
	AATTTCCCTT	TTACTCAGGA	GTTACCCCTC	TTTGATATTT	TCTGTAATTC	TTCTAGACT	GGGGATACAC	8750
	CGTCTCTTGA	CATATTACAC	GTTTCTGTGA	CCACCTGTTA	TCCCATGGGA	CCCACTGCGA	GGGCAGCTGG	8820
55	GAGGCTGCAG	GCTTCAGGTC	CCAGTGGGCT	TGCCATCTGC	CAGTAGAAAC	CTGATGTAGA	ATCAGGGCGC	8890
	AAGTGTGGAC	ACTGTCTCGA	ATCTCAATGT	CTCAGTGTGT	GCTGAAACAT	GTAGAAATTA	AAGTCCATCC	8960
	CTCTACTCT	ACTGGGATTG	AGCCCCCTCC	CTATCCCCC	CCAGGGGCGA	AGGAGTTCCT	CTCACTCCTG	9030
	TGGAGGAAGG	AATGATACTT	TGTTATTTTT	CAGTGTCTGG	ACTGAATCCA	CTGTTTCATT	TGTTGGTTTG	9100
	TTTGTTTTGT	TTTGAGAGGC	GGTTTCACTC	TGTTGTCTCA	GGCTGGAGGG	AGTGAATAGG	CGCGATCTTG	9170
60	GCTTACTGCA	GCCTCTGCCT	CCCAGGTTC	AGTGATTCTC	CTGCTTCCGC	CTCCCATTTG	CGTGGATTGA	9240
	CAGGCACCCG	CCACCATGCC	CAGCTAATTT	TTTGTATTTT	TAGTAGAGAC	GCGGGTGGGT	GGGGTTCACC	9310
	ATGTTGGCCA	GGCTGGTCTC	GAATCTCTGA	CCTCAGATGA	TCCACCTGCC	TCTGCCTCCT	AAAGTGTCTG	9380
	GATTACAGGT	GTGAGCCACC	ATGCCACGCT	CAGAATTTAC	TCTGTTTAGA	AACATCTGGG	CTGAGGTAG	9450
	GAAGCTCACC	CAACTCAAGT	GTTGTGGTGT	TTTAAGCCAA	TGATAGAATT	TTTTTATTGT	TGTTAGAAC	9520
65	CTCTTGATGT	TTTACACTGT	GATGACTAAG	ACATCATCAG	CTTTTCAAAG	ACACACTAAC	TGCACCCATA	9590
	ATACTGGGGT	GTCTTCTGGG	TATCAGCAAT	CTTCATTGAA	TGCCGGGAGG	CGTTTCTCTG	CCATGCACAT	9660
	GGTGTTAATT	ACTCCAGCAT	AATCTCTGCT	TTCATTTTCT	TCTCTTCCCT	CTTTTAAAA	TGTGTTTTCT	9730
	ATGTTGGCTT	CTCTGCAGAG	AACCACTGTA	AGCTACAAC	TAACTTTTGT	TGGAACAAAT	TTTCCAAACC	9800
	GCCCTTTTGC	CCTAGTGGCA	GAGACAATTC	ACAAACACAG	CCCTTTAAAA	AGGCTTAGGG	ATCACTAAGG	9870
70	GGATTTCTAG	AAGAGCGACC	TGTAATCTTA	AGTATTTACA	AGACGAGGCT	AACCTCCAGC	GAGCGTGACA	9940
	CGCCAGGGAG	GGTGCAGGGC	CTGTTCAAAT	GCTAGCTCCA	TAAATAAGC	AATTTCTCC	GGCAGTTTCT	10010
	GAAAGTAGGA	AAGGTTACAT	TAAAGGTTGC	GTTTGTAGC	ATTTCTAGT	TTGCCGACCT	CAGCTACAGC	10080
	ATCCCTGCAA	GGCCTCGGGA	GACCCAGAAG	TTTCTCGCCC	CCTTAGATCC	AAACTTGAGC	AACCCGAGT	10150
	CTGGATTCTT	GGGAAGTCTT	CAGCTGTCTT	GCGGTGTGTC	CGGGGCCCCA	GGTCTGGAGG	GGACAGTGG	10220
75	CCGTGTGGCT	TCTACTGCTG	GGCTGGAAGT	CGGCTCTCCT	AGCTCTCGAC	TCCGAGGCTT	GGAGCCAGGT	10290
	GCCTGGACCC	CGAGGCTGCC	CTCCACCTCT	TGCGGGCGGG	ATGTGACCAG	ATGTTGGGCT	CATCTGCCAG	10360
	ACAGAGTGCC	GGGGCCGAGG	GTCAGGCGCG	TTGTGGCTGG	TGTGAGGCGC	CCGGTGCGCG	CCGAGGAGGA	10430
	GCGCTGGCT	CCATTTCCTCA	CCCTTTCTCG	ACGGGACCGC	CCCGTGGGTT	GATTAACAGA	TTTGGGGTGG	10500

5 TTTGCTCATG GTGGGGACCC CTCGCCGECT GAGAACCTGC AAAGAGAAAT GACGGGCGTG TGTCAAGGAG 10570  
 CCCAAGTCGC GGGGAAGTGT TGCAGGAGAG CACTCCGGGA GGTCCCGCGT GCCCGTCCAG GGAGCAATGC 10640  
 GTCCTCGGGT TCGTCCCGAG CCGCGTCTAC GCGCTCCGT CCTCCCTTC ACCTCCGGCA TTCTGGTGC 10710  
 CCGGAGCCCG ACGCCCCCGG TCCGGACCTG GAGGAGCCCG TGGGTCTCCG GATCAGGCCA GCGGCCAAAG 10780  
 GGTGCGCCGA CGCACCTGTT CCCAGGGCCT CCACATCATG GCCCTCCCT CGGGTTACCC CACAGCCTAG 10850  
 GCCGATTCTGA CCTCTCTCCG CTGGGGCCCT CGCTGGCGTC CCTGCACCT GGGAGCGCGA GCGGCGCGCG 10920  
 GCGGGGGAAG CGCGGCCAG ACCCCCGGGT CCGCCCGGAG CAGCTGCGCT GTCGGGGCCA GCGCGGGCTC 10990  
 CCAGTGGATT CGCGGGCACA GACGCCCAGG ACCCGCTCC CCACGTGGCG GAGGACTGG GGACCCGGGC 11060  
 10 ACCCGTCTCT CCCCTTCACC TTCCAGCTCC GCCTCCTCCG CCGCGACCCC GCCCGTCCC GACCCCTCCC 11130  
 GGGTCCCGGG CCCAGCCCCC TCCGGGCCCT CCCAGCCCT CCTCTTCCT TCCGCGGCC CCCTCTCTCC 11200  
 TCGCGCGCGG AGTTTCAGCG AGCGCTGCGT CCTGCTGCG ACCTGGGAAG CCCTGGCCCC GGCCACCCCT 11270  
 GCGATGCGCG CGCTCCCGG CTGCCGAGCC GTGCGCTCCC TGCTGCGCAG CCCTACCCG GAGGTGCTGC 11340  
 CGCTGGCCAC GTTCGTGCGG CGCTGGGGG CCCAGGGCTG GCGGTGGTG CAGCGCGGGG ACCCGGGCGG 11410  
 15 TTTCGCGCGG CTGGTGGCCC AGTGCTGGT GTGCGTGCCC TGGGACGCAC GGCGCCCCC CGCCGCCCCC 11480  
 TCCTTCCGCC AGGTGGGCCCT CCCCGGGGTC GCGCTCCGCG TGGGGTTAG GGCGCGCGG GGAACCCAG 11550  
 GACATGCGGA GAGCAGCGCA GCGGACTCAG GCGGCTTCCC CCGCAGGTGT CCTGCCTGAA CGGCTCTGGG 11620  
 GCCCGAGTGC TGACAGAGCT GTGCGAGCGC GCGCGGAAGA ACCTGCTGCG CTTCGGGCTC GCGCTGCTGG 11690  
 ACGGGGCCCG CGGGGGCCCC CCCGAGCCCT TCACCAACAG CGTGCGCAGC TACCTGCCCA ACACGGTGAC 11760  
 CGACGCACGT CGGGGAGCGG GGGCGTGGGG GCTGCTGCTG CGCCCGCTGG GCGACGACGT GCTGGTTCAC 11830  
 20 CTGCTGGCAC GCTGCGCGCT CTTTGTGCTG GTGGCTCCCA GCTGCGCTA CCAGGTGTGC GGGCGCGCG 11900  
 TGTACACAGT CGCGCTGCCC ACTCAGGCCG GCGCCCCCGC ACACGCTAGT GGACCCCGAA CGGCTCTGGG 11970  
 ATGCGAACGG GCCTGGAACC ATAGCGTCAG GGAGGCCGGG GTCCCTCTGG GCCTGCCAGC CCCGGGTGCG 12040  
 AGGAGGCGCG GGGGCACTGC CAGCCGAAGT CTGCGGTTGC CCAAGAGGCC CAGGCGTGGC GCTGCCCTG 12110  
 AGCCGGAGCG GACGCCCGTT GGGCAGGGGT CCTGGGCCCA CCCGGGCGAG ACCTGTGGAC GAGTGTACCG 12180  
 25 TGGTTTCTGT GTGGTGTAC CTGCCAGACC CGCCGAAGAA GCCACCTCTT TGGAGGTGCG GCTCTCTGGC 12250  
 ACCGCCCACT CCCACCCATC CGTGGGCCG CAGCACCAG CAGGCCCCCC ATCCACATCG CGGCCACCAC 12320  
 GTCCCTGGGA CACGCCCTGT CCCCGGGTGT ACGCCGAGAC CAAGCACTTC CTCTACTCT CAGGCGACAA 12390  
 GGAGGAGCTG CGGCCCTCTC TCCTACTCAG CTCTCTGAGG CCCAGCTGA CTGGCGCTCG GAGGCTCTG 12460  
 30 GAGACCATCT TTCTGGGTTG CAGGCCCTGG ATGCGAGGGA CTCCCGCAG GTTGCCCGCG CTGCCCGAG 12530  
 GCTACTGGCA AATGCGGCC CTGTTTCTGG AGCTGCTTGG GAACACGCG CAGTGCCCTC ACGGGGTGCT 12600  
 CCTCAAGACG CACTGCCCCG TSCGAGTGC GGTACCCCCA GAGCCGGTG TCTGTGCCG GGAAGAGCCC 12670  
 CAGGGCTCTG TGGCGGCCCC CGAGGAGGAG GACACAGACC CCCGTGCGCT GGTGACAGT CTCCGCCAG 12740  
 ACAGCAGCCC CTGGCAGGTG TACGGGCTTG TCGGGGCGTG CTGCGCGCG CTGGTGCCCC CAGGCTCTG 12810  
 35 GGGCTCCAGG CACAACGAAC GCGGCTTCTC CAGGAACACC AAGAAGTTCA TCTCCCTGGG GAAGCATGCC 12880  
 AAGCTCTCGC TCGAGGAGCT GACGTGGAAG GGGACTGCGG TTGGCTGCGG TGGGCGCCAG 12950  
 GTGAGGAGGT GGTGGCGCTC GAGGGCCAG GCCCGAGAG TGAATGCAGT AGGGGCTCAG AAGAGGGGG 13020  
 AGGCAGAGCC CTGGTCTCTC TGCTCCATC GTACGCTGGG CACACGTGGC TTTTCGCTCA GGACGCTGAG 13090  
 TGGACACGGT GATCTCTGCC CTCTGCTCCA GTTTGCATAA ACTTACGAGG TTCACCTTCA T13160  
 40 CGTTTGTATG GACACCGCGT TTCCAGGCGC CGAGGCCAGA GCAGTGAACA GAGGAGGCTG GCGCGCGGAG 13230  
 TGGAGCCCGG TTGCCGGCAA TGGGGAGAAG TGTCTGGAAG CACAGACGCT CTGGCGAGGG TGCTGCAAG 13300  
 TTACCTATAA TCCTCTTCGC AATTTCAAGG GTGGGAATGA GAGGTGGGGA CGAGAACCCC CTCTTCTTGG 13370  
 GGGTGGGAGG TAAGGGTTT GCAGGTGCAC GTGGTCAGCC AATATGCAGG TTTGTGTTA AGATTTAAT 13440  
 GTGTGTTGAC GCGCAGGTGC GGTGGCTCAC GCGGTAATC CCAGCACTT GGGAGAGCTGA GGCAGGTGGA 13510  
 45 TCACCTGAGG TCAGGAGTTT GAGACAGGCC TGACCAACAT GGTGAACCC TATCTGTACT AAAATACAA 13580  
 AATATTAGCT GGCATGGTGG TGTGTGCTG TAATCCCAGC TACTTGGGAG GCTGAGGCG GAGAATCACT 13650  
 TGAACCCAGG AGCGCGGAGG TGCACTGAGC TGAGATTGTG CCATTGTACT CCAGCCTGGG CGACAAGAGT 13720  
 GAAACTCTGT CTTTAAAAAA AAAAAGTGTG CGTTGATTGT GCCAGACAG GGTAGAGGGA GGGAGTAAAG 13790  
 ACTGTTCTCC AGCAGAGATC CTGGTCCCAT CTTTAGGTAT GAAGAGGGCC ACATGGGAGC AGAGGACAGC 13860  
 50 AGATGGCTCC ACCTGCTGAG GAAGGGACAG TGTTTGTGGG TGTTCAGGGG ATGGTGCTGC TGGGCCCTGC 13930  
 CGTGTCGCCA CCCTGTTTTT CTGGATTGTA TGTGAGGAA CCTCCGCTCC AGCCCCCTT TGGCTGCCAG 14000  
 TGCTCCAGG CCCTACCGTG GCAGCTAGAA GAAGTCCCGA TTTCACCCCC TCCCCACAAA CTCCCAAGAC 14070  
 ATGTAAGACT TCCGGCCATG CAGACAAGGA GGGTGACCTT CTTGGGGCTC TTTTTTTCT TTTTCTCTT 14140  
 TTATGGTGGG AAAAGTCATA TAACATGAGA TTGGCACTCC TAACACCGTT TTCTGTGTAC AGTGCAAGAT 14210  
 55 TGCTAACTCG GCGGTGTTTA CAGCAGGTG CTTGAAATGC TGCGTCTTGC GTGACTGGAA GTCCCTACCC 14280  
 ATCGAACCGG AGCTGCCCTCA CACCTGCTGC GGCTCAGGTG GACCACGCGG AGTCAGATAA GCGTCATGCA 14350  
 ACCCAGTTTT GCTTTTGTG CTCCAGCTTC CTTCTGTGAG GAGAGTTTGA GTTCTCTGAT CAGGACTCTG 14420  
 CCTGTCAATT CTGTTCTCTG ACTTCAGATG AGGTACAAAT CTGCCCCTG CTTATGCAAG GAGTGAGGCG 14490  
 TGGTCCCCCG GTGTCCCTGT CACGTGCAGG GTGAGTGAGG CGTGGCCCC AGGTGTCCCT GTGACGTGTA 14560  
 60 GGGTGAGTGA GGC CGCGGCC CGGGGTGTC CTGTCCGCTG CAGCGTGATT GAGGTGTGCG CCCCGGTTG 14630  
 CCCTGTACAG TGTAGGGTGA GTGAGGCGCC ATCCCCGGT GTCCCTGTCA CGTGTAGGGT GAGTGAGGCG 14700  
 TGGTCCCCCG GTGTCCCTGT CCCGTGAGG GTGAGTGAGG CACTGTCCCC GGGTGTCCCT GTACGTGCA 14770  
 GGGTGAGTGA GGC CGCGGTCC CGGGGTGTC CTCTCAGGTG TAGGGTGAGT GAGGCGCGG CCCAGGGTGT 14840  
 CCCTGTACAG TGTAGGGTGA GTGAGGCACC GTCCCTGGT GTCCCTCCCA GGTATAGGGT GAGTGAGGCA 14910  
 65 CTGTCCCCCG GTGTCCCTGT CACGTGCAGG GTGAGTGAGG CGCGGCCCCC GGGTGTCCCT CTCAGGTGCA 14980  
 GGGTGAGTGA GGC GCTGTCC CTGGGTGTC CTGTCTCGTG TAGGGTGAGT GAGGCTCTGT CCCAGGTGT 15050  
 CCTTGGCGTT TGCTCACTTG AGCTTGCTCC TGAATGTTG CTCTTTCTAT AGCCACAGCT CGCGCGGTTG 15120  
 CCCATTGCCT GGGTAGATGG TGCAGGCGCA GTGCTGGTCC CCAAGCCTAT CTTTCTGAT GCTCGGCTCT 15190  
 TCTTGGTAC CTCTCCGTTT CATTGTTGTA CGGGGACAG GAGCTGCAG CTCTCGCTC CGCGGTGCA 15260  
 70 GGCAGTGCAG CCACAGCTTC AGGTCCGCTT GCCTCTGTTG GGCCTGGCT GCTCACCAG TGCCCGCCAC 15330  
 ATGCATGCTG CCAATACTCC TCTCCAGCT TGCTCATGCG CGAGGCTGGA CTCTGGGCTG CTGTGTCTG 15400  
 CTGCGACGTC TTGCTGAGA CATCCAGAA AGGGTTCTCT GTGCCCTGAA GGAAAGCAA TCACCCAGC 15470  
 CCGCTCACTT GTCTGTTTT CTCCCAAGCT GCGCTCTGC TTGGCCCCCT TGGGTGGGT GCAACGCTG 15540  
 TCACCTTATT CTGGGACCTT GCGCTCATT GCTTAGGCTG GGCTCTGCT CCAGTCCGCC CCTCACATG 15610  
 75 ATTGACGTCC AGCCACAGGT TGGAGTGTCT CTTGCTGTC GTGCTCTGA GACCCACGT GAGGCGCGG 15680  
 GTCTCGGCA GCCTTCGTCA GACTTCCCTC TTGGGCTTA GTTTTGAATT TCACTGATT ACCTCTGAGC 15750  
 TTTCTATCTC TCCATTGTAT GCTTTTCTT GTTTTATCT TTCTAGCTT TTAGTTTGT 15820  
 CATGCCCTTC CCTCTAAGT GTGCCCTTACC TGCACCTGT GTTTTGTAT GAAGTAATCT CAACATCAGC 15890

	CACTTTCAAG	TGTTCTTAAA	ATACCTTCAA	GTGTTAATAC	TTCTTTTAAG	TATTCTTATT	CTGTGATTTT	15960
	TTTCTTTTGT	CACGCTGTGT	TTTGACGTGA	AATCATTTTG	ATATCAGTGA	CTTTTAAGTA	TTCTTTAGCT	16030
	TATTCTGTGA	TTTCTTTGAG	CAGTGAGTGA	TTTGAACACT	GTTCATGTTC	AAGATATGTA	GAGTATCAAG	16100
	ATACGTAGAG	TATTTTAAGT	TATCATTTTA	TTATTGATTT	CTAACTCAGT	TGTGTAGTGG	TCTGTATAAT	16170
5	ACCAATTATT	TGAAGTTTGC	GGAGCCTTGC	TTTGTGATCT	AGTGTGTGCA	TGGTTTCCAG	AACTGTCCAT	16240
	TGTAAATTTG	ACATCCTGTC	AATAGTGGGC	ATGCATGTTT	ACTATATCCA	GCTTATTAAG	GTCCAGTGCA	16310
	AAGCTTCTGT	CTCCTCTAG	ATGCATGAAA	TTCCAAGAAG	GAGGCCATAG	TCCCTCACCT	GGGGGATGGG	16380
	TCTGTTCATT	TCTTCTCGTT	TGGTAGCATT	TATGTGAGGC	ATTGTTAGGT	GCATGCACGT	GGTAGAATTT	16450
10	TTATCTTCCT	GATGAGTGAA	TCTTTTGGAG	ACTTCTATGT	CTCTAGTAAT	CTAGTAATTC	TTTTTTTAAA	16520
	TGTCTCTTAG	TACTGCCACA	CTGGGCTTCT	TTTGATTAGT	ATTTCTCTGC	TGTGTCTGTT	TTCTGCCTTT	16590
	AAITTTATATA	TATATATATA	TTTTTTTTTT	TTTTGAGACA	GAGTCTTGGT	CTGTGCGCCA	GGGTGAGTGC	16660
	AGTGGTGTGA	TCACAGGTCA	GTGTAACCTT	TACCTTCTGG	CCTGAGCCGT	CCTCTCACCT	CAGCCTCCTG	16730
	AGTAGCTGGA	ACTGCAGACA	CGCACCCGTA	CACCTGGCTA	ATTTTTAAAT	TTTTTCTGGA	GACAGGGTCT	16800
15	TGCTGTGTTG	CCCAGGCTGG	TCTCAAACCT	TGGAGTCAA	GGGATCCATC	TACCTCGGCT	TCCCAAAGTG	16870
	CTGAATTACA	GGCATGAGCC	ACCATGTCTG	GCCTAAATTT	CAACACTTTT	ATATTCTTAT	AGTGTGGGTA	16940
	TGTCCTGTGA	ACAGCATGTA	GGTGAATTTT	CAATCCAGTC	TGACAGTCGT	TGTTTAACTG	GATAACCTGT	17010
	TTTATTTTCA	TTTTTTTGTG	ACTAGAGACC	CGCCTGGTGC	ACTCTGATTC	TCCACTTGCC	TGTTGATGTT	17080
	CCTCGTTCCC	TTGTTTCTCA	CCACCTCTTG	GGTTGCCATG	TGCGTTTCTT	GCCGAGTGTG	TGTTGATCCT	17150
20	CTCGTTGCC	CCTGGTCACT	GGGCATTGCT	TTTTATTCTT	CTTTGCTTAG	TGTTACCCCT	TGATCTTTTT	17220
	ATTGTGCTTG	TTTGTCTTGT	TTTATTGAGA	CAGTCTCACT	CTGTCAACCA	GGCTGGAGTG	TAATGGGACA	17290
	ATCTCGGCTC	ACTGCAACCT	CTGCCTCTCT	GGTTCACAGC	GTTCCTCATC	CTCAACCTCA	TGAGTAGCTG	17360
	GGATTACAGG	CGCCACCCAC	CACGCTTGGC	TAATTTTTGT	ATTTTGTAGT	GAGATAGGCT	TTCACCATGT	17430
	TGGCCAGGCT	GGTCTCAAAC	TCCGTACCTC	AAGTGATCTG	CCCGCCTTGG	CCTCCACACG	CACTGGGATG	17500
25	ACAGGTGCAA	GCCACCGTGC	CGGCATACCT	TTGATCTTTT	AAAATGAAGT	CTGAACATT	GCTACCTTGT	17570
	TCCTGAGCAA	TAAGACCTTT	AGTGTATTTT	AGCTCTGGCC	ACCCCCAGC	CTGTGTGCTG	TTTTCCCTGC	17640
	TGACTTAGTT	CTATCTCAGG	CATCTTGACA	CCCCCAACAG	CTAAGCATTA	TTAATATTGT	TTTCCGTGTT	17710
	GAGTGTCTCT	GTAGCTTTGC	CCCCGCCCTG	CTTTTCTCTC	TTTGTCTCCC	GTCTGTCTTC	TGCTCTCAGG	17780
	CCGCCGCTCT	GGGTCCCTCT	CCTTGTCTCT	TGCGTGGTTC	TTCTGTCTTG	TTATTGCTGG	TAAACCCACG	17850
30	CTTTACCTGT	GCTGGCTCTC	ATGGCATCTA	GCGACGCTCG	GGGACCTCTG	CTTATGATGC	ACAATGAGT	17920
	ATGTGGAGAC	TCACGAGGAG	GGCGGTCATC	TTGGCCCGTG	AGTGTCTGGA	GCACCCAGTG	GCCAGCGTTC	17990
	CTTAGCCAGT	GAGTGACAGC	AACGCTCCGT	CGGCCTGGGT	TCAGCTCGGA	AAACCCACAG	CATGTCGGGG	18060
	TCTGGTGGCT	CCGCGGTGTC	GAGTTTGAAG	TCGCGCAAC	CTGCGGTGTC	GCGCCAGCTC	TGACGGGTGT	18130
	GCCTGGCGGG	GGAGTGTCTG	CTTCTCTCTT	TCTGCTTGGG	AACACAGGAC	AAGGATGAGG	CTCCGAGCCG	18200
35	TTGTGCGCCA	ACAGGAGCAT	GACGTGAGCC	ATGTGGATAA	TTTTAAATTT	TCTAGGCTGG	GCGCGGTGGC	18270
	TCACGCTGTG	AATCCACAGC	CTTTGGGAGG	CCAAGCGGGG	TGGATCACGA	GGTCAGGAGG	TCGAGACCAT	18340
	CCTGGCCAAC	ATGATGAAC	CCCATCTGTA	CTAAAAAC	AAAAATTAGC	TGGCGCTGGT	GGCGGGTGCC	18410
	TGTAATCCCA	GCTACTCGGG	AGGCTGAGGC	AGGAGAATTG	CTTGAACCTG	GGAGTTGGAA	GTTCGACGTA	18480
	GCCGACATTG	CACCACTGCA	CTCCAGCCTG	GCAACACAGC	GAGACTCTGT	CTCAAAAAAA	AAAAAAAAAA	18550
40	AAAAAATAAA	AATTTCTAGT	GCCACATTAA	AAAGTAAAAA	AAGAAAAAGT	GAAATTAATG	TAATAATAGA	18620
	TTTTACTGAA	GCCACGATG	TCCACACCTC	ATCATTTTAG	GGTGTATTG	GTGGAGCAT	CACCTACAGG	18690
	ACATTTGACA	TTTTTTGAGC	TTTGTCTGCG	GGATCCCGTG	TGTAGGTCCT	GTGCGTGGCC	ATCTCGGCCT	18760
	GGACCTGCTG	GGCTTCCCAT	GGCCATGGCT	GTGTACCATG	ATGGTGACAG	TCCGGGATGA	GGTCCGCAGG	18830
	CCCTGGATGA	GCTGGATGTG	CAGTGTCCGG	ATGGTGACAG	TCTGGGATGA	GGTCGCCAGG	CCCTGCTGTG	18900
45	AGCTGGATGT	GTGGTGTCTG	GATGGTGCAG	GTCAGGGGTG	AGGTCTCCAG	GCCCTCGGTT	AGCTGGAGGT	18970
	ATGGAGTCCG	GATGATGCAG	GTCGGGGGTG	AGGTCTCCAG	GCCCTGCTGT	GAGCTGGATG	TGTGGTGTCT	19040
	GGATGGTGCA	GCTCAGGGGT	GAGGTCTCCA	GGCCCTCGGT	AAGCTGGAGG	TATGGAGTCC	GGATGATGCA	19110
	GGTCCGGGGT	GAGGTGCGCA	GGCCCTGCTG	TGAGCTGGAT	GTGTGGTGTG	TGGATGGTGT	AGGTCTGGGG	19180
	TGAGGTCACC	AGGCCCTGCG	GTGAGCTGGG	TGTGCGGTGT	CTGGATGGTG	CAGGTCTGGA	GTGAGGTCCT	19250
50	CAGACGGTGC	CAGACCATGC	GGTGAGCTGG	ATATCGCGTG	TCCGGATGGT	GCAGGTCTGG	GCTGAGGTTG	19320
	CCAGGCCCTG	CTGTGAGTTG	GATGTGGGGT	GTCCGGATGC	TGCAGGTCCG	GTGTGAGGTC	ACCAAGGCCCT	19390
	GCTGTGAGCT	GGATGTGTGG	TGTCTGGATG	GTGCAAGTCT	GGGGTGAAGG	TGCGCAGGCC	CCTGCTTGTG	19460
	AGCTGGATGT	GTGGTGTCTG	GATGGTGCAG	GTCTGGAGTG	AGGTGCGCAG	GCCCTCGGTT	CCTGAGTGTG	19530
	GCAGTGTCCA	GATGGTGCAG	GTCGGGGGTG	AGGTGCGCAG	ACCCTGCGGT	GAGCTGGATG	TCCGGTGTCT	19600
	GGATGGTGCA	GGTCTGGAGT	GAGGTGCGCA	GGCCCTCGGT	GAGCTGGATG	TATGGAGTCC	GGATGGTGCC	19670
55	GGTCCGGGGT	GAGGTGCGCA	GACCCCTGCT	TGAGCTGGAT	GTGCGGTGTC	TGGATGGTAC	AGGTCTGGAG	19740
	TGAGGTGCGC	AGACCCCTGT	GTGAGCTGGA	TATCGGTTGT	CCGGATGGTG	CAGGTGAGGG	GTGAGGTCTC	19810
	CAGGCCCTCG	GTGAGCTGGA	GGTATGGAAT	CCGGATGATG	CAGGTGCGGG	GTGAGGTCTC	CAGGCCCTGC	19880
	TGTGAACCTG	ATGTGCGGCG	TCTGGATGGT	GCAGTCTCTG	GCTGTGCTCG	CCAGGCCCTC	GGTGAAGCTG	19950
60	AGGTATGGAG	TCCGGATGAT	GCAGGTCCGG	GGTGAGGTCT	CCAGGCCCTG	CTGTGAGCTG	GATGTGCGGC	20020
	GTCTGGATGG	TGCAGGTCTG	GGGTGTGGTC	GCCAGGCCCT	CGGTGAGCTG	GAGGTATGGA	GTCCGGATGA	20090
	TGCAGGTCCG	GGGTGAGGTT	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCTG	TATCCGGATG	GTGCAAGTCC	20160
	GGGTGAGGTC	GCCAGGCCCT	GCTGTGAGCT	TATCCGGATG	GTGCAAGTCT	GGGTGAGGTC	GGGTGAGGTC	20230
	CACCGAGCCC	TGCGGTGAGC	TGGTTGTGCG	GTGTCCGGTT	GCTGCAAGTC	CGGGGTGAGT	TCGCGAGGCC	20300
65	CTCGGTGAGC	TGGATGTGCG	GTGTCCCGCT	GTCCGGATGG	TGCAGGTCCA	GGGTGAGGTC	GCTAGGCCCT	20370
	TGGTGGGCTG	GATGTGCCGT	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGAAGCT	20440
	GATGTGCCGT	GTCTGCATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	GATGTGTGGT	20510
	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	GCTGTGAGCT	GGATGTGCGG	TGTCTGGATG	20580
	GTGCAAGTCC	GGGTGAGGTC	AGCCAAAGGCC	TTCCGTGAGC	TGGATGTGGG	GTGTCCGGAT	GGGTGAGGTC	20650
70	CGGGGTGAGG	TGCGCAAGGCC	CTCGGTTAG	CTGGATATGC	GGTGTCCGGA	TGGTGCAGGT	CCGGGTGAGT	20720
	GTCACCAAGC	CCTGCGGTTA	GCTGGATGTG	CGGTGTCTGG	ATGGTGCAGG	TCCGGGGTGA	GGTCCCAAGG	20790
	CCCTGCTGTG	AGCTGGATGT	GCTGTATCCG	GATGGTGCAG	GTCCGGGGTG	AGGTGCGGCT	GGCCTGCAGT	20860
	GAGCTGGATG	TGCTGTATCC	GGATGGTGCA	GGTCTGGCGT	GAGGTGCGCA	GGCCCTGCGG	TTAGCTGGAT	20930
	ATGCGGTGTC	GGATGGTGCA	GGTCCGGGGT	GAGGTCAACA	GGCCCTGCGG	TTAGCTGGAT	GTGCGGTGTC	21000
	CGGATGGTGC	AGGTCTGGGG	TGAGGTGCGC	GAGGTGCGCA	GGTGTGCGG	TGTGCTGTAT	CCGGATGGTG	21070
75	CAGGTCCGGG	GTGAGGTGCG	CAGGCCCTGC	GGTGAAGTGG	ATGTGCTGTA	TCCGGATGGT	GCAGGTCTGG	21140
	CGTGAGGTGC	CCAGGCCCTG	CGGTGAGCTG	GATGTGAGT	GTACGGATGG	TGCAGGTCCG	GGGTGAGGTC	21210
	GCCAGGCCCT	CGGGTGGGCT	GTATGTGTGT	TGCTGTGATG	GTGCAAGTCC	GGGGTGAAGT	CGCCAGGCC	21280

	TGCGGTGAGC	TGGATGTGTG	GTGTCTGGAT	GCTGCAGGTC	CGGGGTGAGT	TCGCCAGGCC	CTCGGTGAGC	21350
	TGGATATCGG	GTGTCCCGGT	GTCCGAATGG	TGCAGGTCCA	GGGTGAGGTC	GCCAGGCCCT	TGGTGGGCTG	21420
	GATGTGCCGT	GTCCGGATGG	TGCAGGTCTG	GGGTGAGGTC	GCCAGGCCCT	TGGTGAAGTC	GATGTGCCGT	21490
	GTCCGGATGG	TGCAGGTCCG	GGGTGAGGTC	ACCAGGCCCT	CGGTGATCTG	GATGTGGCAT	GTCTCTCTCG	21560
5	TTTAAGGGGT	TGGCTGTGTT	CCGCCCGCAG	AGCACCGTCT	GGGTGAGGAG	ATCCTGGGCA	AGTTCCTGCA	21630
	CTGGCTGATG	AGTGTGTACG	TCGTGAGGCT	GCTCAGGTCT	TTCTTTTATG	TCACGGAGAC	CACGTTTCAA	21700
	AAGAACAGCC	TCTTTTCTA	CCGGAAGAGT	GTCTGGAGCA	AGTTGCAAA	CATTGGAATC	AGGTACTGTA	21770
	TCCCCAGGCC	AGGCCCTCTG	TTCTCGAAGT	CCTGGAACAC	CAGCCCGGCC	TCAGCATGCG	CCTGTCTCAA	21840
	CTTGCCCTGTG	CTTCCCTGGC	TGTGAGGCTC	TGGGTGCGGA	GCCAGGGGCC	CCGTACAGAG	CCTGGTCCAA	21910
10	GTGGATTCTG	TGCAAGGCTC	TGACTGCGTC	GAGCTCACGT	TCTCTTACTT	GTAATAACAG	GAGTTTGTGC	21980
	CAAGTGGTCT	CTAGGGTTTG	TAAAGCAGAA	GGGATTATAA	TTAGATGGAA	ACACTACCAC	TAGCCTCCTT	22050
	GCCTTTCCCT	GGGATGTGGG	TCTGATTCTC	TCTCTCTTTT	TTTTTCTTTT	TTTGAGATGG	AGTCTCACTC	22120
	TGTTGCCAG	GCTGGAGTGC	AGTGGCATAA	TCTTGGCTCA	CTGCAACCTC	CACCTCCTGG	GTTTAAGCGA	22190
	TTTACCAGCC	TCAGGCTCCT	AAGTAGCTGG	GATTACAGGC	ACCTGCCACC	ACGCTTGCTG	AATTTTGTGA	22260
15	CTTTTAGGAG	AGACGGGGTT	TCACCATGTT	GGCCAGGCTG	GTCTCGAACT	CATGACCTCA	GGTGATCCAC	22330
	CCACCTTGGC	CTCCCAAGT	GCTGGGTTTA	CAGGCTAAGC	CACCGTGCCC	AGCCCCCGAT	TCTCTTTTAA	22400
	TTTCATGCTG	TCTGTATGAA	TCTTCAATCT	ATTGGATTTA	GGTCATGAGA	GGATAAAATC	CCACCCACTT	22470
	GGCGACTCAC	TGCAGGGAGC	ACCTGTGACG	GGAGCACCTG	GGGATAGGAG	AGTTCCACCA	TAGCCTTAAGT	22540
20	TCTAGGTGGC	TGCATTGAA	TGGCTGTGAG	ATTTGTGCTG	CAATGTTCCG	CTGATGAGAG	TGTGAGATTG	22610
	TGACAGATTG	AAGCTGGATT	TGCATCAGTG	AGGGACGGGA	GCGCTGGTCT	GGGAGATGCC	AGCCTGGCTG	22680
	AGCCAGGGCC	ATGGTATTAG	CTTCTCCGTG	TCCCGCCACG	GCTGACTGTG	GAGGCTTTTA	GTCAGAAGAT	22750
	CAGGGCTTCC	CCAGCTCCCC	TGCACACTCG	AGTCCCTGGG	GGGCTTGTG	ACACCCCATG	CCCCAAATAC	22820
	GGATGTCTCG	AGAGGGAGCT	GGCAGCAGAC	CTCGTCAGAG	GTAACACAGC	CTCTGGGCTG	GGGACCCCGA	22890
	CGTGGTGCTG	GGGCCATTTC	CTTGATCTCG	GGGGAGGGTC	AGGGCTTTCC	CTGTGGGAAC	AAGTTAATAC	22960
25	ACAAATGACC	TTACTTAGAC	TTTACACGTA	TTTAATGGTG	TGCGACCCAA	CATGGTCATT	TGACCATGAT	23030
	TTTGGAAAGA	ATTTAATTGG	GGTGACCGGA	AGGAGCAGAC	AGACGTGGTG	GTCCCAAGA	TGCTCCTTGT	23100
	CACACTGTTG	ACTGTTGTTT	TGCTGGGGG	CCCTTGGAGG	CCCCCTCTCC	CTGGACAGGG	TACCGTGCCT	23170
	TTTCTACTCT	GCTGGGCTG	CGGCTGCGG	TGAGGGCACC	AGCTCCGGAG	CACCCGCGGC	CCAGTGTGCC	23240
	ACGGAGTGCC	AGGCTGTGAG	CCACAGATGC	CCAGGTCCAG	GTGTGGGCGC	TCCAGCCCCC	TGCCCCCAT	23310
30	GGGTGCTTTT	GGGGGAAAG	GGCAAGGGCA	GAGCTGTGAG	GAGACTGGTG	GGCTCATGAG	AGCTGATTCT	23380
	GCTCCTTGGC	TGAGCTGCCC	TGAGCAGCCT	CTCCCGCCCT	CTCCATCTGA	AGGGATGTGG	CTCTTCTTAC	23450
	CTGGGGGTCC	TGCTTGGGGC	CAGCCTTGGG	CTACCCCACT	GGCTGTACCA	GAGGGACAGG	CATCCTGTGT	23520
	GGAGGGGCAT	GGGTTACAGT	GGCCCCAGAT	GCAGCCTGGG	ACCAGGCTCC	CTGGTGCTGA	TGGTGGGACA	23590
35	GTCACCTTGG	GGGTGACCCG	CCGAGCTGGG	CGTCCCCCAG	GTTGACTATA	GGACAGAGTG	TCCAGGTGCC	23660
	CTGCAAGTAG	AGGGGCTCTC	AGAGGCGTCT	GGCTGGCATG	GGTGGACGTG	GCCCCGGGCA	TGACCTTCTG	23730
	CGTGTGCTGC	CGTGGGTGCC	CTGAGCCCTC	ACTGAGTCCG	TGGGGGCTTG	TGGCTTCCCG	TGAGCTTCCC	23800
	CCTATCTGTG	TGTCTGGGCT	AGCAAGCCTC	CTGAGGGGCT	CTCTATTGCA	GACAGCACTT	GAGAGGGGTG	23870
	CAGCTGCGGG	AGCTGTGCGA	AGCAGAGGTC	AGGCAGCATC	GGGAAGCCAG	GCCCCCGCTG	CTGACGTCCA	23940
40	GACTCCGCTT	CATCCCCAAG	CCTGACGGGC	TGCGGCCGAT	TGTGAACATG	GACTACGCTG	TGGGAGGCCA	24010
	AACGTTCCCG	AGAGAAAAGA	GGGTGGCTGT	GCTTTGGTTT	AACTTCTCTT	TTAAACAGAA	TGCGCTTTGA	24080
	GGCCCACTT	TGGTATCAGC	TTAGATGAAG	GGCCCGGAGG	AGGGGCCACG	GGACACAGCC	AGGGCCATGG	24150
	CACGGCGCCA	ACCCATTGTG	GGCAGCAGTG	AGGTGGCCGA	GGTGCCGGTG	CCTCCAGAAA	AGCAGCGTGG	24220
	GGGTGTAGGG	GGAGCTCCTG	GGGCAGGGAC	AGGCTCTGAG	GACCACAAGA	AGCAGCCGGG	CCAGGGCCTG	24290
45	GATGCAGCAC	GGCCCCAGGT	CCTGGATCCG	TGCTCTGCTG	TGGTGCGCAG	CCTCCGTGCG	CTTCCGCTTA	24360
	CGGGGCCCCG	GGAGCAGGCC	ACGACTTGCA	GGAGCCCCAC	GGGCTCTGAG	GATCCTGGAG	CTTCCGCTTA	24430
	GGCTCTGCA	CCCCACCCCT	GTGGCTGCGG	TGGCTGCGGT	GACCCCGTCA	TCTGAGGAGA	GTGTGGGGTG	24500
	AGGTGGAGAC	AGGTGTGGCA	TGAGGATCCC	GTGTGCAACA	CACATGCGCG	CAGGAACCCG	TTTCAAAACG	24570
	GGTCTGAGGA	AGCTGGGAGG	GGTTCTAGGT	CCCGGGTCTG	GGTGGCTGGG	GACACTGGGG	AGGGGCTGCT	24640
50	TCTCCCTTGG	GTCCCTATGG	TGGGGTGGGC	ACTTGCCCGG	ATCCACTTTC	CTGACTGTCT	CCCATGCTGT	24710
	CCCCGCCAGG	CCGAGCGTCT	CACCTCGAGG	GTGAAGGCAC	TGTTACAGCGT	GCTCAACTAC	GAGCGGGCGC	24780
	GGCGCCCCGG	CCTCTTGGGC	GCCTCTGTGC	TGGGCCGTGA	CGATATCCAC	AGGGCCTGGC	GCACCTTCTG	24850
	GCTCGGTGTG	CGGGCCCGAG	ACCCGCCCGC	TGAGCTGTAC	TTTGTCAAGG	TGGGTGCCGG	GGACCCCCGT	24920
	GAGCAGCCCT	GCTGGACCTT	GGGAGTGCGT	GCCTGATTGG	CACCTCATGT	TGGGTGGAGG	AGGTACTCCT	24990
55	GGGTGGGCGG	CAGGGAGTGC	AGGTGACCCCT	GTCACGTGTT	AGGACACACC	TGGCACCTAG	GGTGGAGGCC	25060
	TTACAGCCTT	CCTGCAGCAC	ATGGGGCCGA	CTGTGCACCC	TGACTGCCCG	GGCTCCTATT	CCCAAGGAGG	25130
	GTCGCCACTG	ATTCCAGTTT	CCGTGAGAGA	AGGAACCCGA	ACGGCTCAGC	CACCAAGGCC	CGGTGCTCTG	25200
	CACCCCAAGT	CTGAGCCAGG	GGTCTCTGTG	CCTGAGGCTC	AGAGAGGGGA	CACAGCCCCG	CTTCCCTTGG	25270
	GGGTCTGGAG	TGGTGGGGGT	CAGAGAGAGA	GTGGGGGACA	CCGCCAGGCC	AGGCCCTGAG	GGCAGAGGTT	25340
60	ATGTCTGAGT	TTCTGCGTGG	CCACTGTGAG	TCTCCTCGCC	TCCACTCACA	CAGGTGGATG	TGACGGGCGC	25410
	GTACGACACC	ATCCCCCAGG	ACAGGCTCAC	GGAGGTCATC	GCCAGCATCA	TCAAAACCCA	GAACACGTAC	25480
	TGCGTGGCTG	GGTATGCCGT	GGTCCAGAAG	CGCCGCCCATG	GGCACGTCCT	CAAGGCCCTT	AAGAGCCACG	25550
	TAAGGTTTCA	GTGTGATAGT	CGTGTCCAGG	ATGTGTGCTC	CTGGGATATG	AATGTGCTTA	GAATGCAGTC	25620
	GTGTCTGTGA	TGCGTTTCTG	TGGTGGAGGT	ACTTCCATGA	TTTACACATC	TGTGATATGC	GTGTGTGGCA	25690
65	CGTGTGTGTC	GTGGTGCAATG	TATCTGTGGC	GTGCATATTT	GTGGTGTGTG	TGTGTGTGGC	ACGTGTGTGT	25760
	CCATGGGTGTG	TGTGCTGTGT	GTGTGCATGT	GTGTGTGTCT	GTGACAGTGT	CATGTTTCATG	CTGTGTGCTG	25830
	CATGTCTGTG	ATGTGCCTAT	TTGTGGTGTG	TGTGTGCATG	TGTCCGTGAC	ATATGCGTGT	CTATGGCATG	25900
	GGTGTGTGTC	GGCCCTTGGC	CTTACTCCTT	CCTCCTCCAG	GCATGGTCCG	CACCATTTGT	CTCAGCTCTT	25970
	CGGGTGCTGG	TTTGGGGAGC	TCCACATTTA	GGGTCTCTAC	TTCTAGCATG	GGTGGCCCTG	TCTTGTGACA	26040
70	GGGCTTGGCC	TTGGAGACTG	TAAGCCAGGT	TTGAGAGGAG	AGTAGGGATG	CTGGTGGTAC	CTTCTTGGAC	26110
	CCCTGGCACC	CCCAGGACCC	CAGTCTGGCC	TATGCCGGCT	CCATGAGATA	TAGGAAGGCT	GATTCAGGCC	26180
	TGCTTCCCGG	GGACACACTC	CTCCAGAGC	GGCCGGGGGC	CTTGGGGCTC	GGCAGGGGTG	AAAGGGGGCT	26250
	TGGGCTTGGG	TTCCACCCCA	GTGGCTCATG	GCACGCTGGA	GGGGTAAGCC	CTCAAAGTCG	TGCCAGGGCC	26320
	GGGTGCAGAG	GTGAAGAAGT	ATCCCTGGAG	CTTCGGTCTG	GGGAGAGGCA	CATGTGGAAG	CCCACAAGGA	26390
75	CCTCTTCTCT	TGACTTCTTG	AGCT					26414

## Contig 2:

	TGTGGGATTG	GTTTTTCATGT	GTGGGATAGG	TGGGGATCTG	TGGGATTGGT	TTTTATGAGT	GGGGTAACAC	70
	AGAGTTCAGG	GCGAGCTTTC	TTCTCTGAGT	GGGTCTGCAG	GTGCTCCAAC	AGCTTTTATTG	AGGAGACCAT	140
5	ATCTTCCTTT	GAACATAGGT	CGGGTTTATA	GTAAGTCAGG	GGTGTGGAGG	CCTCCCTCTGG	GCTCCCTGTT	210
	CTGTTTCTTC	CACCTCTGGG	TCGTGTGGTG	CCTGCTGTGG	TGTGTGGCCG	GTGGGCAGGG	CTTCCAGGCC	280
	TCCTTGTGTT	CATTGGCCCTG	GATGTGGCCC	TGGCTACGCT	CCGTCTCTGG	AATTCCCTCTG	CGAGTTGGAG	350
	GCTTTCTTTT	TTTTCTTTTT	TCTTTTTTTT	TTTTTTTTTT	TGATAACAGA	GTCTCGCTCT	TTTTTGCCCA	420
10	GGCTGGAGTG	GTTTGGCGTG	ATCTTGGCTC	ACTGCAACCT	GTGCTTCCTG	AGTTCAGCA	ATTCCTCTGC	490
	CTCAGCCTCC	CAAGTAGCTG	GAATTATAGG	CGCCACCCAC	CATGCTGACT	AATTTTTGTA	ATTTTAGTAG	560
	AGACGAGGTT	TCTCCATGTT	GGCCAGGCTG	GTCTCGAACT	CCTGACCTCA	GGTGATCCTC	CCACCTCGGC	630
	CTCCCAAAGT	GCTGGGATGA	CAGGTGTGAA	CCGCCGCGCC	CGGCCGAGAC	TCGCTTCCTG	CAGCTTCCTG	700
	GAGATCTGCA	GCGATAGCTG	CCTGCAGCCT	TGGTGCTGAC	AACCTCCGTT	TTCTTCTCC	AGGTCTCGCT	770
15	AGGGGTCTTT	CCATTTCATG	ACTCTCTTCA	CAGAAGAGTT	TCACGTGTGC	TGATTTCCCG	GCTGTTTCTT	840
	GCCTAATTGG	TGCTGTCTGT	TTATCGATGG	CCTCCTTCCA	TTTCTTTAG	GCTTTGTTTA	TGTTGTGTTT	910
	TCCGGCTCCT	TGAAGGAAAA	GTTTCGATTA	TGGATGTTTG	AACCTTCTTT	TCTAAACAAG	CATCTGAAGT	980
	TGCCGTTTTT	CCTCTAAAGC	AGGGATCCCG	AGGCCCTTGG	CTGTGGAGTG	GCACCGGTCT	GGGGCTGTTT	1050
	AGGAACCCGG	GCGACAGCGG	GAGGCTAGGT	GGGGTGTGGG	GAGCCAGCGT	TCCCGCTTGA	CGCCCGCCCC	1120
20	TCTCAGATCA	GCAGTGGCAT	CGGGTGCTCA	GAGGCGCACA	CACCCCTACTG	AGAAGTGTGC	GTGAGAGGGG	1190
	TCTCAGATTCT	GTGCTCCTTA	TGGGAATCTA	ATGCCCTGATG	ATCTGAGGTG	GAACCGTTTG	CTCCCAAAAC	1260
	CATCCCCCTC	CCCACTGTCTG	TCCTGTGGAA	AAATCGTCTT	CCACGAAACC	AGTCCCTGGT	ACCACAATGG	1330
	TTGGGGACCC	TGTGCTAAAG	ACCTGCTTCA	GCAGCCTCTC	GTCACTGTTG	ATATATTGGC	TTTTCTGTGT	1400
	TGAGTCCAGA	ATAATTACGG	ATTTCTGTGA	TGCTTTCGGC	CGACCTCAGA	CCCATGGGCT	ATTTGTGGGC	1470
25	GTGTTGCTTG	CTCCTGGGTT	GGGAAGGGTG	CAGGCCCATC	GTACCTTCTCT	GTTACTGCCT	TCCAGGTGGG	1540
	TTCTCAGGGT	TGAATCGTAC	TCGATGTGGT	TTTAGCCAC	GGCCCTGCGG	CCAGCTCCTG	GGGGCTGGGG	1610
	AACATGCTGA	AGCACAGAGT	CACCGTGCGC	GTCTTTTGAT	GCCTCACAAG	CTCAGAGGCT	CCTGTGTCCG	1680
	TGTTAGTGTG	TGTACGTGTC	CTGCTCACAT	CCTGTCTTGG	GGACGACGGG	GCTTAGCAGG	TCCCGTAGTA	1750
	AATGACAAGC	GTCTTGGGGG	AGTCTGCAGA	ATAGGAGGTG	GGGGTGGCGG	TCTCTCTCCC	GCCTCTTCAG	1820
30	ACTCTTCTCC	TGCCTGTGCT	GTGGCTGCAC	CTGCATCCCT	GCAATCCCTC	CAGCACTGGG	CTGGAGAGGC	1890
	CCGGGAGCTC	GAGTGCCACT	TGTGCCACGT	GACTGTGGAT	GGCAGTCGGT	CACGGGGGTC	TGATGTGTGG	1960
	TGACTGTGGA	TGGCGGTTGG	TCACAGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2030
	TGTGTGACT	GTGGATGGCG	GTGCTGGGGT	CTGATGTGTG	GTGACTGTGG	ATGGCGGTCTG	TGGGGTCTGA	2100
	TGTGGTGA	GTGGATGGCG	GTGCTGGGGT	CTGATGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2170
35	TGGTGA	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2240
	TGGTGA	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2310
	TGGTGA	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2380
	TGGTGA	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2450
	TGGTGA	GGATGGCAGT	CGTGGGGTCT	GATGTGTGGT	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	2520
40	GGTCTGATGT	GTGGTGA	TGGATGGTGA	TCGGTCAACG	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	2590
	GTGCTGGGGT	GTGACTGTGG	ATGGCGGTTG	GTCCCGGGGG	TCTGATGTGT	GGTGA	GTGCTGTG	2660
	GATGGCGATC	GGTCAACAGG	GTCTGATGTG	TGGTGA	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2730
	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	TGGTGA	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2800
	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	TGGTGA	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2870
45	GACTGTGGAT	GGCGGTCTGTG	GGGTCTGATG	TGGTGA	GGATGGCGGT	CGTGGGGTCT	GATGTGTGGT	2940
	GTGGTGA	TGGATGGCAG	TCGTGGGGTCT	TGATGTGTGG	TGACTGTGGA	TGGCGGTCTGT	GGGTCTGAT	3010
	GTGGTGA	TGGATGGCAG	TCGTGGGGTCT	TGATGTGTGG	TGACTGTGGA	TGGCGGTCTGT	GGGTCTGAT	3080
	ATGTTGTTGG	ACTGTGGATG	GGGTCTGTGG	GGTCTGATGT	GGTGA	GATGGCGGTCT	GTGGGGTCTG	3150
	ATGTTGTTGG	ACTGTGGATG	GGGTCTGTGG	GGTCTGATGT	GGTGA	GATGGCGGTCT	GTGGGGTCTG	3220
50	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3290
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3360
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3430
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3500
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3570
55	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3640
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3710
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3780
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3850
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3920
60	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	3990
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4060
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4130
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4200
65	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4270
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4340
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4410
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4480
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4550
70	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4620
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4690
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4760
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4830
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4900
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	4970
	GGGTCTGATG	TGTGGTGA	GTGGATGGCG	GTGCTAGGGT	CTGATGTGGT	GTGACTGTGG	ATGGCAGTCC	5040

	CTGCCCGGGG	CCACCAGAGT	CTCCTTTCT	GGCCCCCGCC	CCCTCCGGCT	CCTGGGCTGC	AGGCTCCCGA	5110
	GGCCCCGGAA	ACATGGCTCG	GCTTGGCGCA	GGCCGAGCGG	AGCAGGTGCC	ACACGAGGCC	TGGAAATGGC	5180
	AAGCGGGGTG	TGGAGTTGCT	CCTGGCTGGA	GGACGAGGGG	CGGGGGGTGT	GTCTGGGTCA	GGTGTGGGCC	5250
5	GAGCGTTTGA	GCCTGCAGCT	TGTCACTCC	AAGTTACTAC	TGACGCTGGA	CACCCGGCTC	TCACACGCTT	5320
	GTATCTCTCT	CTCCCGATAC	AAAAGGATTT	TATCCGATTC	TCATTCCTGT	CCCTGTCTGT	TGACCCCGCG	5390
	GAGGGCGCGG	GCTCTTCTCT	CTGTGACTAG	ATTTCCCATC	TGGAAAGTGC	GGGGTTGACC	GTGTAGTTTG	5460
	CTCCTCTCGG	GGGGCCTGTG	GTGGCCATGG	GGCAGGCGGC	CTGGGAGAGC	TGCCGTGACA	CAGCCACTGG	5530
	GTGAGCCACA	CTCACGGTGG	TAGAGCCACA	TTGCCCTGGT	CCACATCAGC	TCCTCTGGAT	TTTAAGTAAA	5600
10	ACCACACACC	TCCCGGCGAG	CATCTGCGTG	CGACCCTGTG	TGTGCCCTGG	GAGAGTGGTA	GCACGGAGGA	5670
	AATTGCTGCA	CACTCAAGGT	CATCAGCAAG	GTCACTCCGA	GTCAGGTGGA	ACGTGGAGGC	CTCTCTCTGG	5740
	GATCGTCTCC	AGCGGATAAA	GGACTGTGCA	CAGCTTCGGA	AGCTTTTATT	TAAAAATATA	ACTATTAAAT	5810
	ATTGCATTAT	AAGTAATCAC	TAATGGTATC	AGCAATTATA	ATATTTATTA	AAGTATAATT	AGAAATATTA	5880
	AGTAGTACAC	ACGTTCTGGA	AAAACACAAA	TTGCACATGG	CAGCAGAGTG	AATTTTGGCC	GAGGACACGC	5950
15	TGTGCACATG	TGTGTAAGCG	GGCCCCAGGC	CCACAGAATT	CGCTGACAAA	GTCACCTCCC	CAGAGAAGCC	6020
	ACCACGGGCC	TCCTTCGTGG	TCGTGAATTT	TATTAAGATG	GATCAAGTCA	CGTACCGTCC	ACGTGTGGCA	6090
	GGGCTTTGGG	GAATGTGAGG	TGATGACTGC	GTCTCATGCG	CCTGACAGAC	AGGAGGTGAC	TGTGTCTGTC	6160
	CTGTCCCTAG	GACACGGACA	GGCCCGAAGC	TCTAGTCCCC	ATCGTGGTCC	AGTTTGGCCT	CTGAATAAAA	6230
	AGTAGTCTAA	AACCTGTTGC	CCCAAAACT	AAGAACAGAG	AGAGTTTCCC	ATCCCATGTG	CTCACAGGGG	6300
20	CGTATCTGCT	TGCGTTGACT	CGCTGGGCTG	GGCCGACTCC	TAGAGTTGGT	GCCTGTGCTT	CTGTGCAAAA	6370
	AGTGCACTCC	TCCTGGCCAT	CACGTGTGATA	TCTGCACCAG	CAAGGAAAGC	CTCTTTTCTT	TCCTTTCTTT	6440
	TTTTTTTTTT	GAGACGGAAC	GTCACTGTGT	TCTGCCTGGG	CTTGAGTGCA	GTGGCGCGAT	CTCAACTCAC	6510
	TGCAACCTCC	GCCTCCCGGG	TTCCAGCATT	TCTCTGCCT	CAGCCTCCCG	AGCAGCTGAG	ATTACAGGCA	6580
25	CCCACCCCTC	GCGCTGGCT	AATTTTGTGA	TTTTTAGTAG	AGAGGGGTTC	TTGCCATGTT	GGCCAGGGCT	6650
	GTCTCGAAGT	CCTGACCTCA	GGTGATCCAC	CCACCTCGGC	CTCCCAAGT	GCTGGGATTA	CAGGTGTGAG	6720
	CCATCAGCCC	CAGCCGGAAA	GCCCTCTTTT	AAGGTGACCA	CCTATAGCGC	TTCGCCAAAA	TAACAGGTCT	6790
	TGTTTTTGCA	GTAGGCTGCA	AGCCTCTCTT	AGCAACAGGA	GTGGCTCTCT	GTGGGCTCTG	GGGATGGCTG	6860
	AGGGTCGCGT	GGCAGCCATG	CCTCTCTGTG	GCACCTTTAG	GTTCACCGGG	GCTATTCTGC	TCTCACTGTT	6930
30	TGTCTGAAAA	CGCACCCCTG	GCATCCTTGT	TTGGAGAGTT	TCTGCTTCTC	GTGTGTCATG	CTGAAACTAG	7000
	GGGCAAGGTT	GTATCCGTTG	GGCGGCAGCG	GCTACATGTA	GGGTCTATGAG	TCTTTCACCG	TGGACAAATT	7070
	CCTTGAAAAA	AAAAAAGGAA	GCTCGGTTAA	GCATTATTC	CGGGTCAAGT	GTCTGGTTCT	GTGAATAAAC	7140
	CTTAAGATT	AAGAAACCTT	AATGAAAGAA	AACCTTGATG	ATTCAAGACA	AGGATGTGGT	CACACCTGTG	7210
	GCTGGATCTG	TTTCAGCCGC	CCCAGTGCA	GGTGAGAGTG	GGGAGCAGGG	ATTGTTTGT	CAGAGGCTCT	7280
35	ATCTGGTATG	TTTCTGAGGT	GTTCGCCGGC	TGAATGGTAG	ACGTGTCGTT	TGTGTGTATG	AGGTTCTGTG	7350
	TCTGTGTGTG	GCTCGGTTTG	AGTGTACGCA	GTGCCAGCAC	ATGCCCTGCC	CGTCTCTCAC	CTGTGTCTTC	7420
	CCGCGCCAGG	TCCTACGTCC	AGTGCCAGGG	GATCCCGCAG	GGCTCCATCC	TCTCCACGCT	CTCAGGCTGC	7490
	CTGTGCTACG	GGCAGATGGA	GAACAAGCTG	TTTGGCGGGA	TTCCGGCGGA	CGGGTGAGGC	CTCCTCTTCC	7560
	CCAGGGGGGG	TTGGGTGGGG	GTGATTTTGC	TTTGATGCA	TTCACTGTGA	ATATTCCTGG	TGCTCTGGAG	7630
	ACCATGACTG	CTCTGTCTTG	AGGAACAGAA	CAAGGTTGCA	GGCCCTTCTT	GGTATGAAGC	CGCACGGGAG	7700
40	GGGTTGCACA	GCCTGAGGAC	TGCGGGCTCC	ACGCAGGCTC	TGTCCAGCGG	CCATGTCCAG	AGGCCCTCAGG	7770
	GCTCAGCAGG	CGGGAGGGCC	GCTGCCCTGC	ATGATGAGCA	TGTGAATTCA	ACACCCAGGA	AGCACACGAG	7840
	CTTCTGTAC	GTCACCCAGG	TTCCGTTAGG	GTCCCTGGGG	AGATGGGGGT	GGTGACAGCT	GAGGGCCCCAC	7910
	ATCTCCAGCG	AGGCCCTCGA	CAGGTGGGCT	GGACTGGGCG	CCTCTTCAGC	CATTGCCCCA	TCCCACTTGC	7980
	ATGGGGTCTA	CACCCAAGGA	CGCACACACC	TAAATATCGT	GCAACCTTAA	TGTGGTTCAA	CTCAGCTGGC	8050
45	TTTTATTGAC	AGCAGTTACT	TTTTTTTTTT	TAATACTTTA	AGTTCTAGGG	TACATGTGCA	CGACGTGCAG	8120
	GTTAGTTACA	TATGTATACA	TGTGCCATGT	TGGTGTGCTG	CACCCATTAA	CTCATCATTT	ACATTAGGTA	8190
	TATCTCTTAA	TGCTATCCCT	CCCCACTCCC	CCCATCCCAT	GACAGGCCCT	GGTGTGTGAT	GTTCCTCCAC	8260
	CTGTGTCCAA	GTGTTCTCAT	TGTTCAAGTC	CCACCTGTGA	GTGAGAACAT	GTGGTGTGTT	GTTTCTTTTC	8330
	CTTGCAATAG	TTTGCTCAGA	GTGATGGTTT	CCAGCTTCGT	CCATGTCCCT	ACAAAGGACA	TGAATCTCAT	8400
50	CTTTTTATAG	ACTGCATAGT	ATTCGCTGGT	GTATATGTGC	CACATTTTCT	TAATCCAGTC	TATCATCGAT	8470
	GGACATTTGG	GTGTTGTGCA	AGTCTTTGCT	ACTGTGAATA	GTGCCGCAAT	AAACATACGT	GTGCATGTGT	8540
	CTTTATAGCA	GCATGATTTA	TAATCCITTT	GGTATATACC	CAGTAATGGG	ATGGCTGGGT	CAAAATGGTAT	8610
	TTCTAGTTCT	AGATCCTTGA	GGAATCACCA	CACGTCTTTC	CACAATGGTT	GAACATAGTT	ACACTCCACAC	8680
	CAACAGTGTA	AAAGTGTTCT	GGTGTGGGAG	AGGATGTGGA	CAGCAGTTAT	TTTTTTATGA	AAATAGTATC	8750
55	ACTGAACAAG	CAGACAGTTA	GTGAAGGATG	CGTCAGGAAG	CCTGCAGGCC	ACACAGCCAT	TTCTCTCGAA	8820
	GACTCCGGGT	TTTTCTGTG	CATCTTTTGA	AACTCTAGCT	CCAATTATAG	CATGTACAGT	GGATCAAGGT	8890
	TCTTCTTCAT	TAAGGTTCAA	GTCTAGATT	GAATAAGTT	TATGTAACAG	AAACAAAAAT	TTCTTGTACA	8960
	CACAACCTGC	TCTGGGATTT	GGAGGAAAGT	GTCCCTGAGC	TGGCGGCACA	CTGGTCAGCC	CTCTGGGACA	9030
	GGATACCTCT	GGCCCATGGT	GATGGGGCGC	TGGGCTTGGG	CCTGAGGGTC	ACACAGTGCA	CCATGCCGAC	9100
60	CTTCTGTGTC	ATAGGATCTG	GGTCTCGGAT	CATGCTGAGG	ACCACAGCTG	CCATGCTGGT	AAAGGGCACCC	9170
	ACGTGGCTCA	GAGGGGGCGA	GGTCCGAGC	CCCAGCTTTC	TTACCGTCTT	CAGTTATTTT	TCCCTAAGAG	9240
	TCTGAGAAAT	GGGGCCGCGC	CTGATGGCCT	TCGTTCTCTC	TCAGCTGGCA	CAGAATTGCA	CAAGCTGATG	9310
	GTAACACTGT	AGTACTTATA	ATGAATGAGG	AATTGCTGTA	GCAGTTAACT	GTAGAGAGCT	CGTCTGTTGG	9380
	AAAGAAATTT	AAGTTTTTCA	TTTAACCGCT	TTGGAGAATG	TTACTTTTAT	TATGGCTGTG	TAAATTTGTT	9450
65	GACATTCAGT	CCCTCGTAGA	CAGATACTAC	GTAAAAAGTG	TAAAGTTAAC	CTTGCTGTGT	ATTTTCCCTT	9520
	ATTTTAGGCT	GCTCCTGCGT	TTGGTGGATG	ATTTCTTGTT	GGTGACACCT	CACCTCACCC	ACGCGAAACG	9590
	CTTCTCAGG	TGAGGCCCGT	GCCGTGTGTC	TGTGGGGACC	TCCACAGCCT	GTGGGCTTTG	CAGTTGAGCC	9660
	CCCGTGTGCC	TGCCCTGTGC	ACCGCAGCGT	TGTCTCTGCC	AAGTCTCTCT	TCTCTGCCGG	TGCTGGATCC	9730
	GCAAGAGCAG	AGGCGCTTGG	CCGTGCACCC	AGGCTTGGGG	GCGCAGGGGC	ACCTTCGGGA	GGGAGTGGGT	9800
70	ACCGTGCAGG	CCCTGGTCCCT	GCAGAGACGC	ACCCAGGTTA	CACACGTGGT	GAGTGCAGGC	GGTGACCTGG	9870
	CTCCTGCTGC	TCTTTGAAAA	GTCAAGAGTG	GCGGCTCTCT	GGGCCCCAGT	GAGACCCCCA	GGAGCTGTGC	9940
	ACAGGGCCCTG	CAGGGCCGAG	GCGGACAGCT	CTCCCCAGG	GTGCACCTGA	GCCTGCGGAG	AGCAGGAGCT	10010
	GCTGAGTGAG	CTGGCCACACA	GCGTTCGCTG	CGGTACAGTT	CCTGCGTGGG	GTGTTTGGG	ATCGGTGGGA	10080
	GAATTTGGAT	TTGCTGAGTG	CTGCTGTCTT	GAACACCGGA	GATGGCTAGG	AGTGGGTTTC	AGAGTTGATT	10150
75	TTTGTGAATC	AAACTAAAA	CAGGCACAGG	GAACCTGGCC	TCAGCACAGG	GGATTGTCCA	ATGTGGTCCC	10220
	CCTCAAGGGC	GCCCCACAGA	GGCGGTGGGC	TTGTTTTAAA	GTGCGATTTG	ACGAGGGGAC	AGAACTCTTG	10290
	AAAGCTGTAA	AGGGAACCTT	CAGAAAAATG	GGCCGCCAGG	GGTGGTTTCA	GGTGCTTTGC	TGGGCTGTGT	10360
	TTGTGAAAC	CCATTGGGAC	CCGCCCTCCA	AGTCCACCCT	CCAGGTCCAC	CCTCCAGGGC	CGCCCTGGGC	10430

	TGGGGGTATG	CCTGGCGTTC	CTTGTGCCGC	AGCCCGGAGC	ACAGCAGGCT	GTGCACATTT	AAATCCACTA	10500
	AGATTCACTC	GGGGGGAGCC	CAGGTCCCAA	GCAACTGAGG	GCTCAGGAGT	CCTGAGGCTG	CTGAGGGGAC	10570
	AGAGCAGACG	GGGAACGCTG	CTTCTGTGTG	GCAAGTTCCT	GAGGGTGCTG	GCCAGGGAGG	TGGCTCAGAG	10640
	TGTATGTGGG	GGTCCCACCG	GGGGCAGAAC	TCTGTCTCTG	ATGAGTCCGC	AGCCATGTAA	CAGGAAGGGG	10710
5	TGGCCACAGG	GAGCTGGGAA	TGCACCAGGG	GAGCTGCCCA	GCTGGCCGAG	GTCCCAGGGC	CAGGGCCACAG	10780
	GAAGGGCAGG	GGGACGCCCG	GGGCCACAGC	AGAGGCCGCA	GGAAGGGGAG	GGGATGCCCA	GGCCAGAGCA	10850
	GAGGCTACCG	GGCACAGGGG	GGCTCCCTGA	GCTGGGTGAG	CGAGGCTCAT	GACTCGGCGA	GGGAACCTCC	10920
	TTGACGTGAA	GCTGACGACT	GGTGTGCCCC	AGCTCACAGC	CCAGCCAGGT	CCCGCGCTG	AGCAGGAATC	10990
	CAGAACCCTC	CCCTTTGTCT	AAAGCACAGC	AGATGCCTTC	AGGGCATCTA	GGAGAAAACA	GGCAAAGTCG	11060
10	TTGAGAAACG	TCTTAAAGA	AGGTGGGATG	GTGGCAATTT	CTTGTCAGG	TTTGTAGCTG	CCCCGGACCA	11130
	CAGATGAGTC	TATAACGGGA	TTGTGGTGT	GCCATGGGGA	CACATGAGAT	GGACCATCAC	AGAGGCCACT	11200
	GGGGCTGCAC	CTCCCATCTG	AGTCTCTGGT	GTCCCGGGTC	CAGGCCAGGT	TCTTGCAATG	TCACCTACCT	11270
	GTCCCTGCCG	GGAGACAGGG	AAAGCACCCC	GAAGTCTGGA	GCAGGGCTGG	GTCCAGGCTC	CTCAGAGCTC	11340
	CTGCCAGGCC	CAGCACCCCTG	CTCCAAATCA	CCACTTCTCT	GGGGTTTTCC	AAAGCATTTA	ACAAGGGTGT	11410
15	CAGGTTACCT	CCTGGGTGAC	GGCCCCGCAT	CCTGGGGCTG	ACATTGCCCC	TCTGCCTTAG	GACCCCTGGT	11480
	CGAGGTGTCC	CTGAGTATGG	CTGCGTGGTG	AACTTGCCGA	AGACAGTGGT	GAACCTTCCCT	GTAGAAGACG	11550
	AGGCCCTGGG	TGGCAGGGCT	TTTGTTCAGA	TGCCGGGCCA	CGGCCTATTG	CCCTGGTGGG	GGCTGCTGCT	11620
	GGATACCCGG	ACCCTGGAGG	TGCAGAGCGA	TACTCTCAGG	TGAGCGCACC	TGGCCGGGAG	GAGAGCCCTG	11690
	CGCCGGCTGG	GGCAGGTGCT	GCTGCAGGGC	CGTTCGCTCC	ACCTCTGCTT	CCGTGTGGGG	CAGGCGACTG	11760
20	CCAATCCCAA	AGGGTCAGAG	GCCACAGGGT	GCCCTCGTGC	CCATCTGGGG	CTGAGCAGAA	ATGCATCTTT	11830
	CTGTGGGAGT	GAGGGTGCTC	ACAACGGGAG	CAGTTTTCTG	TGCTATTTTG	GTAAAAGGAA	ATGGTGACCC	11900
	AGACCTGGGT	GCACTGAGGT	GTCTTCAGAA	AGCAGTCTGG	ATCCGAACCC	AAGACGCCCG	GGCCCTGCTG	11970
	GGCGTGAATC	TCTCAAAACC	GAACACAGGG	GCCCTGCTGG	GCATGAGTCC	CTCTGAACCC	TAGAGCCCTG	12040
25	GGCCCTGTGT	GGCGTGAATC	TCTCCGAACC	CAGAGACTTC	AGGGCCCTTT	TGGCGGTGAG	TCTCTCCGCT	12110
	GTGAGCCCCA	CACCTCAAGG	CTCATCCACA	GTCTACAGGA	TGCCATGAGT	TCATGATCAC	GTGTGACCCA	12180
	TCAGGGGACA	GGCCCATGGT	GTGGGGGGGG	TCTCTACAAA	ATTCTGGGGT	CTTGTTTCCC	CAGAGCCCGA	12250
	GAGCTCAAGG	CCCCGTCTCA	GGCTCAGACA	CAATGAATT	GAAGATGGAC	ACAGATGCAG	AAATCTGTGC	12320
	TGTTTCTTTT	ATGAATAAAA	AGTATCAACA	TCCAGGCAGG	GGCAAGGTGG	CTCACACCTA	TAATCCACAG	12390
30	ACTTTGGGAG	GGCGAGGTGG	GTGGATCACT	TGAGGCCAGG	AGTTTGAGGC	CAACCTAACC	AACATAGTGA	12460
	AATTCATTTT	CTACTTAAAA	AATACAAAAA	TAGCCCTGGC	CTGGTGGCAC	ACGCCCTGAT	TCCCGCTAT	12530
	CGCGGAGGCT	GAGGCAGGAG	AATCATTTGA	ACCCAGGAGG	CAGAGGTGGC	AGTGAGCCGA	GATCACACCA	12600
	CTGCACTCCA	GCCTGGGCAA	CAGAGTGAGA	CTTCATCTTA	AAAAAAGTATC	AGCATTTCCA	12670	
	AACCATAGTG	GACAGGTGTT	TTTTTATTCT	GTCCCTCGAT	AATATTACT	GGTGCTGTGC	TAGAGGCCGG	12740
35	AACTGGGGGT	GCCTTCTCT	GAAAGGCACA	CCTTCATGGG	AAGAGAAATA	AGTGGTGAAT	GGTTGTTAAA	12810
	CCAGAGGTTT	AAACTGGGGT	CCTGTCTGTC	TGAGTTAAACA	GTCCAGATCT	GGACTTTTGC	TCTTTCCAGA	12880
	ATGCTCCCTG	GGGTTTGCTT	CATGGGGGAG	CAGCAGGTGT	GGACACCTCT	GTGATGGGGG	AGCAGCAGGT	12950
	CGAGACGCCC	TCATGATGGG	GGAGTGGCAG	GTGCAGACAC	CCTTGTGCAT	GGTGCCGAGC	ATGTCCTTGT	13020
	TGCAGCTCCC	TCCCCACAAG	GATGCCGGTC	TCTGTGCTC	CCCACAGTCC	CTGCTTCCCT	CTCACAGCCT	13090
40	TACCTGGTCC	TGGCCTCCAC	TGGCTTTTGT	TGCATGATT	CCACATTTCC	TGGGCTCCCA	GCACCTCTTC	13160
	GCCTCTCCCA	GGCAGCTCTG	CAGGTGCTGGC	CATACCAGTC	AGCTGTGAAC	TGTCCACTGC	TTATTTGCT	13230
	CCCCATGAAA	TGTATTTTTT	AGGACAGGCA	CCCCGTGGTC	CAGCCTCTGG	CACAGCATCA	GTGAATGTTA	13300
	TTGAAGGACA	AAGGACAGAC	AAACAAATCA	GGAAAATGGG	TTCTCTCTAA	ACACATTGCA	AAGCCACAGA	13370
	GGCTAGTGCA	GGATGGGTGG	GCATCAGGTC	ATCAGATGTG	GGTCCAAATG	CAGAATATTC	TGTCTCCCA	13440
45	AAGGCCACTT	GGTCAGAGTG	TGTGCTTGA	GAGGTGGCTC	TAAAAGTCA	GCAGTGGAGG	CAGTGGTTCG	13510
	CCATACTCAG	GGTGAAGTCA	CATCCTCTGT	GTCTGAAGTA	TACAGCAGAG	GCTTGAAGGG	CATCTGGGAG	13580
	AAGAAAACAG	GCAAAATGAT	TAAGAAAAGT	GAAAAGGAA	AAGTGGTAAG	ATGGGAAATTT	TCTTGTCCAG	13650
	ATTTTAGTCT	CCCAAACCC	AGCTCAGATG	GTAGAAATGT	GTGAGAACTG	ATGGACAGAA	CAATAGAGCA	13720
	AAACGGGAAG	CCTATCTCTC	AGAAACGTGT	GTAAATGTGG	TATGTGGCAC	AGCTGATGGA	AAAGAGAGTG	13790
50	TGTGTGTAAT	TTTTTTTCT	GAGAAAAGT	ACTGGAAAGCA	AATAAGTTGT	GTCTTTACAG	CATATACCAG	13860
	AGCAGATTCT	AGGTAGAAGA	GGAGACACAT	GCAAAACA	CCAGCAACAG	AAATAAAACA	AAAGACTCAA	13930
	AGGGGAAGGA	GGTGAAGCTT	CCCTGGTTTG	GTGTTGGGGA	AGGACACACA	GGGAGGCGGA	TGAAACCACT	14000
	GAGGCAACGG	GCATTGCTTT	CACCTGCAGAG	AAACTCAGCT	TGCCCTGAGCC	ACAGTGAAAA	TGGCCATTCC	14070
	CTGGAGCGTT	TGTGCACGTG	ATTTATTTAA	GGGCCCTCTG	GAGGTCCTGC	ACATTCACTC	TCTCACTTTG	14140
55	TTCTCCTAAC	CACCTGAGAG	GTAGAGGAGG	AAAGGCTCCA	GGGAGCAGC	CGCCCTTGGT	CACCCAGCTG	14210
	GCAAAGGGCA	TGCATGATTG	CAGCCTGGCC	TCCTGCTCCG	GGGCCCTTGC	TCTGCCCGAG	GACCCACAC	14280
	AAGTCAGACC	CATAGGCTCA	GGGTGAGCCG	GAGCCCAAGG	TCGTGTGGG	GATGGCTGTG	AAAGAAGAAA	14350
	TGGAGCTCTG	ATGCACACTT	GGGAAGGTCC	TACCAGCAGC	GTCAAAGAAA	TGCATGTGAA	ACTGACAGCG	14420
	AGACCCATCC	CTCAAAGAAA	CGCACGTGAA	ACTGATGGCG	AGACCTGTCC	CCATCCCTCA	TGCTGGCTCC	14490
60	TTTTCTGGGC	TTGCCAAGAG	CCAGCATCAG	GTGAGGCA	GCTGGAAGA	CTTTCTGGA	AAGCAGCTTG	14560
	TTTGATGGA	AGTCCTCACA	ATGTCCTGTG	TCTTCCAGT	AATTCCACTT	CTGAAGTGAC	CAGACATTAT	14630
	CACGGGTCTT	ATTTACCATT	TCCAGTGTTC	CAGGCAGGGG	GACTTGCCAC	AGCAAGTCAC	GAACCTGCC	14700
	AAATACAGGG	CTAAGGAGAT	ATTATGCATC	ACAAAACCTG	CTCTGCCATT	AAACATTTT	CAAGAATTT	14770
	TTGAAGAATG	TTTAATGGCA	CAAAACGTTT	ATTTCAATGT	AGCAGTGTTC	AAAGCTGGAT	GTAAAAGAAC	14840
65	ACACCCAGG	AGCCTGCCGT	GAATGTGATG	TGTGTTTATC	TTTGGACATG	GACATACATG	GGCAGTGAGT	14910
	GGTGGTGAGG	CCCTGGAGGA	CATCGGTGGG	ATGCCCTCAT	CCTGCCCTCT	TGGAGACACC	ATGTTGCGCA	14980
	CGTGCACTCA	CTGGAGCCCT	GTTTAGCTGG	TGCCACCTGG	CTCTTCCATC	CCTGAGATTC	AAACAGCTG	15050
	AGATTCCCA	CGCCCAACTC	AGTGTCTTCC	CACAAAAAAC	CTGAGTCACA	CCTGTGTTCA	CTCGAGGGAC	15120
	GCCCGGGAGC	CAGGGCTCCA	CAGTTTATTA	TGTGTTTTTG	GCTGAGTTAT	GTGCAGATCT	CATCAGGGCA	15190
70	GATGATGAGT	GCACAAACAC	GGCCGTGCGA	GGTTTGGATA	CACCTCAACAT	CACATAGCCAG	GTCTTGGTGG	15260
	AGTTTGGTCA	TGCAGAGTCT	GGATGGCATG	TAGCATTTGG	AGTCCATGGA	GTGAGCACCC	AGCCCTCTGC	15330
	GGCTCCAGCG	CATGCCCCAG	CGAGGACAAG	GAAGCGGGAG	GGCTCTTTGG	GGCTCTTTGG	AGCAAGCTTT	15400
	CGAGGAGGGG	GCTGGGTGTG	GGGACGGCAC	CTGTGTCTGA	CATCCCCCCC	TGTGTCTCAG	CTATGCCCGG	15470
	ACCTCCATCA	GAGCCAGTCT	CACCTTCAAC	CGCGGCTTCA	AGGCTGGGAG	GAACATGCGT	CGCAAACTCT	15540
75	TTGGGGTCTT	CGCGCTGAAG	TGTCACAGCC	TGTTTCTGGA	TTTGCAGGTT	AGCAGGCTGA	TGGTCAGCAC	15610
	AGAGTTTCAGA	GTTTCAGGAGG	TGTTGTGCGA	AGTATGTGTG	TGTGTGTGTG	CGCGCGTGCC	TGCAAGGCTG	15680
	ATGGTGACTG	GCTGCACGTA	AGAGTGCACA	TGTACGCATA	TACACGTGAG	CACATACATG	TGTGATCTGT	15750
	TGTACATGAA	GGCATGGCAG	TGTGTGCACA	GGTGTGCAAG	GGCACAGATG	TGTGCACATG	CGAATGCACA	15820

	CCTGACATGC	ATGTGTGTTT	GTGCACAGTC	GTGTGGGCAT	TCACGTGAGG	TGCATGCGTG	TGGGTGTGCA	15890
	GTGTGAGTAG	CATGTGTGCA	CATAACATGT	ATTGAGGGGT	CCTCGTGTTT	ACCCCGCTAG	GTCTCAGCA	15960
	CCAGTGCCAC	TCCTTACAGG	ATGAGACGGG	GTCCAGGGC	TTGGTGGGCT	GAGGCTCTGA	AGCTGCAGCC	16030
5	CTGAGGGCAT	TGTCCCATCT	GGGCATCCGC	GTCCACTCCC	TCTCCTGTGG	GCTTCTGTGT	CCACTCCCCC	16100
	TCTCCTGTGG	GCATTACAT	CCACTCCACT	CCCTCTCTCC	TGTGGGCATC	CGCGTCCACT	CCCCCTCTCT	16170
	GTGGGCATCT	CGGTCCACCT	CCCCCTCTCT	TGGGCATTTG	CGTCCACTCC	CTCTCCTGGT	TCCTTCCGTG	16240
	CTTGGCCGAG	CCTCGGGGGC	AGGCAGATGA	CACAGAGTCT	TGACTCGCCC	AGGGTGGTTC	GCAGCTGCCG	16310
	GGTGAGGGCC	AGGCCGATT	TCACTGGGAA	GAGGGATAGT	TTCTTGTCAA	AATGTTCCCT	TTTCTTGTTT	16380
10	CATCTGAATG	GATGATAAAG	CAAAAAGTAA	AAACTTAAAA	TCCAGAGAGG	GTTTCTACCG	TTTCTCACTC	16450
	TTTCTTGGCG	ACTCTAGGTG	AACAGCCTCC	AGACGGTGTG	CACCAACATC	TACAAGATCC	TCCTGTGTGA	16520
	GGCGTACAGG	TGAGCCGCCA	CCAAGGGGTG	CAGGCCACGC	CTCCAGGGAC	CCTCCGCGCT	CTGCTCACTT	16590
	CTGACCCGGG	GCTTCACCTT	GGAACTCCTG	GGTTTATAGG	GCAAGGAATG	TCTTACGTTT	TCAGTGGTGC	16660
	TGCTGCTGTG	GCACAGTTCT	GTTCCGCTGG	CTCTGTGCAA	AGCACCTGTT	CTCCATCTCT	GGGTAGTGGT	16730
15	AGGAGCCGGT	GTGGCCCCAG	GTGTCCCCAC	TGTGCTGTGT	CACTGGCCGT	GGGAGCTCAT	GGAGGCCATC	16800
	CCAGGGCAGC	AGGGGCATGG	GGTAAAGAGA	TGTTTATGGG	GAGTCTTAGC	AGAGGAGGCT	GGGAAGGTGT	16870
	CTGAACAGTA	GATGGGAGAT	CAGATGCCCG	GAGGATTGGG	GGTCTCAGCA	AAGAGGGCCG	AGGTGGGTGC	16940
	AGGTGAGGGT	CGCTGGCCCC	ACCCCGGGGA	AGGTGCAGCA	GAGCTGTGGC	TCCCCACACA	GGCCGGCCAG	17010
20	CACCTGTGCT	CTGGGCATGG	CTGTGCTCCT	GGAACTGTCC	CTGTCTGGCC	TGGTCAAGGG	GTGCCCTCTG	17080
	CAAGAATCGA	CAACTTTATC	ACAGAGGGAA	GGGCCAATCT	GTGGAGGCCA	CAGGGCCAGC	TTCTGCTCTG	17150
	ATGCAGGGCA	GGTGGTGGCA	CAAGCCTCGG	GGCTGTACCA	AAGGGCAGTC	GGGCACCACA	GGCCCGGGCC	17220
	TTCACCTCAA	CAGGCCCTCC	GAGCCACTCG	GAGCTGAATG	CCAGGAGGCC	GAGGCCCTCG	CCCATGTAGG	17290
	GCTGAGAAGG	AGTGTGAGCA	TTTGTGTTAC	CCAGGGCCGA	GGCTGCGCGA	ATTACCGTGC	ACACTTGATG	17360
	TGAAATGAGG	TCGTGCTCTA	TCGTGAAAGC	CCAGCAAGGG	CTCAGGGGAG	AGTTTTCAT	TACAAGGTGC	17430
25	TACCATGAAA	ATGGTTTTTA	ACCCGAGTGC	TTGGCCCTTC	ATGCTCTGGC	AGGGAGGGCA	GAGCCACAGC	17500
	TGCTATGTTT	CGCCTTTGCA	CCAGCTCCAG	AGGCTTGGGA	CCAGGCTGTC	TCAGTTCCAG	GGTGCCTCCG	17570
	GCTCAGACCG	CCCTCCTCTC	TGCTCTCTCT	CTCTGCCCTA	AATCTTCCCT	CGTTTGCATC	TCCCTGAGCC	17640
	GTGCTTGGGC	CCTCGTGCAG	GCTGCTTAC	TCCTTTCCTG	AAACCCCTTG	GGTGTGCTGG	ATACAGGTGC	17710
	CACCTGAGGAC	TGGAGGTGTC	TGACACTGTG	TTGACCCCA	GGGTCCAGCT	GGCGTGTCTG	GGGCTCTCTT	17780
30	GGGCCATGAT	GAGGTTCAGAG	GAGTTTTCCT	AGGTGAAAGC	TCCTGGGAAA	CTCCACGGGC	CATGTGACCT	17850
	GCCACCTGCT	CCTCCCATAT	TCAGCTCAGT	CTTGTCTCTA	TTTCCCCACC	AGGGCTCTCT	GCTCCGAGGA	17920
	GCTCCCTAG	AGGGCCTGGG	CTCAGGGCAG	GGCGGTGAG	TTTCCCCACC	CATGTGGGGA	CCCTTGGGTA	17990
	GTGCTTGTAT	TGGGTAGCCC	TGAGGAGGCC	GAGATGCGAT	GGGCCACGGG	CCGTTTCCAA	ACACAGAGTC	18060
35	AGGCACGTGG	AAGGCCCAGG	AATCCCTTTC	CCTCGAGGCA	GGAGTGGGAG	AACGGAGAGC	TGGGCCCGCA	18130
	TTTCACGGCA	GCCAGGCTGC	AGTGGGCGAG	GCTGTGGTGG	TCCACGTGGC	GCTGGGGGGC	GGGTCTGATT	18200
	CAAAATCCGCT	GGGGCTCGGC	CTTCTGGGCC	CGTCTGGGCC	GGGCCCTCAC	ACGGGCTTGG	GTGGGACGCC	18270
	CCGACCTCTA	GCAGGTGGCT	ATTCTCCCTT	TTGGAAGAGA	GGCCCTCACC	CATGCTAGGT	GTTTCCCTCC	18340
	TGGGTCAAGGA	GGGTGGCCGT	GTGGCAACCC	CGGGACCTTA	GGCTTATTTA	TTTGTTTAAA	AACATTCTCG	18410
40	GCTTGGCTTC	CGTTGTGTCT	AAATGGGGAA	AAGACATCCC	ACCTCAGCAG	AGTTACTGAG	AGGCTGAAAC	18480
	CGGGGTGCTG	GCTTGAAGTG	TGTGATCTCA	GGTCATTCCA	GAAGTGGCTC	AGGAAGTCAG	TGAGACACAG	18550
	TACATGGGGG	GCTCAGGAG	TGGGTGAGAT	GAGGTACACG	GGGGGCTCAG	GCAGTGGGTG	AGGCCAGGTA	18620
	CATGGGGGGC	TCAGGCAGTG	GGTGAGATGA	GGTACACGGG	GGGCTCAGGC	AGAGGGTCAG	ACCAGGTACA	18690
	CGGGGGCTCT	GATCACACGC	ACATATGAGC	ACATGTGCAC	ATGTGCTGTT	TCATGGTAGC	CAGGTCTGTG	18760
45	CACACCTGCC	CCAAAGTCCC	AGGAAGCTGA	GAGGCCAAAG	ATGGAGGCTG	ACAGGGCTGG	CGCGGTGGCT	18830
	CACACCTGTA	GTCCAGCAC	TTTGGGAGGC	CGAGGCCAGA	GGATCCCTTG	AGCCCCAGGAG	TTTAAGACCA	18900
	GCCTGAGCAA	CATAGTAGAA	CCCCATCTCT	ATGAAAAATA	AAAAACAAAA	TTAGCTGAAC	ATGTTGGTGT	18970
	CGCGCTGTAG	TTCCAATACT	TGGGAGGGCT	AAGTGGGAGG	ATCACTTGAG	CCCAGGAGGT	GGAAAGTGA	19040
	GTGAGCTGAG	ATTGCACCA	TGTACTGCAG	CCTGGGTGAC	AGAGTGAGAG	CCCATCTCAA	CAACACCAAA	19110
50	GAAGACTGAC	AAATGCAGTT	TCTTGGAAAG	AAACATTTAG	TAGGAACCTA	ACCTACACAC	AGAAGCCAA	19180
	TCGGTGTCTC	GGTGTGAGTG	AGATGAGATG	ATGGGTCTCT	ACACCATCAC	CCCAGACCCA	GGGTTTATGC	19250
	ACCACAGGGG	CGGGTGGCTC	AGAAAGGATG	CGCAGGACGT	TGATATACGA	TGACATCAAG	GTGTCTGAC	19320
	GAAGGGCAGG	ATTATGATA	AGTACCTGCT	GGTACACAAG	GAACAATGGA	TAAACTGGAA	ACCTTAGAGG	19390
	CCTTCCCGGA	ACAGGGGCTA	ATCAGAAGCC	AGCATGGGGG	GCTGGCATCC	AGGATGGAGC	GTCTTCAGCC	19460
55	TCCACATGCG	TGTTTATACA	GATGGTGCAC	AGAAACGCAG	TGTACCTGTG	CACACACAGA	CACGCAGCTA	19530
	CTCCGACACA	CAAGCACACA	CACAGACATG	CATGCATGCA	TCCGTGTGTG	TGCACCTGTG	CCCATGAGGA	19600
	AACCCATGCA	TGTGCATTCA	TGCACGCACA	CAGGCACCCG	TGGGCCCATG	CCCACACCCA	CGAGCACCGT	19670
	CTGATTAGGA	GGCCTTCTCT	CTGACGCTGT	CCGCCATCCT	CTCAGGTTTC	ACGCATGTGT	GCTGCAGCTC	19740
	CCATTTCATC	AGCAAGTTTG	GAAGAACCCT	ACATTTTCTC	TGCGGCTCAT	CTCTGACACG	GCCTCCCTCT	19810
60	GCTACTCCAT	CCTGAAAGCC	AAGAAGCAG	GTATGTGCAG	GTGCTTGGCC	TCAGTGGCAG	CAGTGCCTGC	19880
	CTGCTGGTGT	TAGTGTGTCA	GGAGACTGAG	TGAATCTGGG	CTTAGGAAGT	TCTTACCCCT	TTTGCATCA	19950
	GGAAAGTGGT	TAACCCAAAC	ACTGTACGGC	TCGTCTGCCC	GGCCTCTCGT	GGGGTGAACA	GAGACCTGTA	20020
	TGGAAGGGAC	AGGAGCTGTG	TGGGAGCTGC	CATCCTTCCC	ACCTTGTCTT	GCCTGGGGAA	CGCCTGGGGG	20090
	GATTTGGGCTG	TCTCCCTGCC	ATGGCACTTA	GGGCCCTTGT	GCAAACCCAG	GCCAAGGGCT	TAGGAGGAGG	20160
65	CCAGGCCCAG	GCTACCCAC	CCCTCTCAGG	AGCAGAGGCC	CGGTATCAC	ACGACAGAGC	CCCGCGCGCT	20230
	CCTCTGCTTC	CCAGTACCCG	TCCTCTGCCC	CTGGACACTT	TGTCAGCAT	CAGGGAGGTT	TCTGATCCGT	20300
	CTGAAATTC	AGCCATGTGC	AACCTGCGGT	CCTGAGCTTA	ACAGCTTCTA	CTTCTGTGTT	TTTCTGTGTT	20370
	GTGGAAATTT	CACCTGGAGA	AGCCGAAGAA	AACATTTCTG	TCGTGACTCC	TGCGGTGCTT	GGGTGGGGAC	20440
70	AGCCAGAGAT	GGAGCCACCC	CGCAGACCGT	CGGGTGTGGG	CAGCTTTCGG	GTGTCTCCTG	GGAGGGGAGC	20510
	TGGGCTGGGC	CTGTGACTCC	TCAGCCTCTG	TTTTCCCCCA	GGGATGTGCG	TGGGGGCCAA	GGGCGCCGCC	20580
	GGCCCTCTGC	CCTCCGAGGC	CGTGCAGTGG	CTGTGCCACC	AAGCATTCCT	GCTCAAGCTG	ACTCGACACC	20650
	GTGTCAACCT	CGTGCCACTC	CTGGGCTCAC	TCAGGACAGG	CAAGTGTGGG	TGGAGGCCAG	TGCGGGCCCT	20720
	ACCTGCCCCAG	GGGTCACTCT	TGAAGCCCTT	GTGTGGGGCG	AGCAGCCTCA	GATGCTGCTG	AAGTGCAGAC	20790
75	GGCCCGGGGC	CTGACCTCTG	GGGCCCTGGG	CCACGCTGGC	AGCCCTATGT	GATTAAACGC	TGGTGTCCCT	20860
	AGGCCACGGA	GCTGGCCAGG	GTCCCAACT	CTGCTTCCCA	TCTCAGGGGC	GATGGCTCCC	TGCGCTCTCC	20930
	CACGCTTGGG	AGCCTTCTGA	CCCTTGACCT	GTGTCTCTCT	ACAGCCTCTT	CCCTGGGTGC	TGCCCTGAGC	21000
	TCCTGGGGTC	CTGAGCAAGT	TCTCTCCCGC	CCCGCCGCGT	CCAGCGTCAC	TGGGCTGCCT	TGCTGCTGCG	21070
	CCCGGTGGAG	GGGTGTCTGT	CCCTTCACTG	AGGTTCCAC	CAGCCAGGGC	CACAGGTGTC	AGGCCCTGCC	21140

TGCCCGGCCA CCCACACGTC CTAGGAGGGT TGGAGGATGC CACCTCTGGC CTCTTCTGGA ACGGAGTCTG 21280  
 ATTTTGGCCC CGCAGCCAG ACGCAGCTGA GTCGGAAGCT CCGGGGAGC ACCTGACTG CCCTGGAGGC 21350  
 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT GGCACCCGC CCACAGCCAG 21420  
 GCGGAGAGCA GACACCAGCA GCCCTGTCAC GCCGGGCTCT GCGGGGAGG GAGGGAGGGG CGGCCACAC 21490  
 CCAGGCCCGC ACGGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCAG GCCTGCATGT CCGGCTGAAG 21560  
 GCTGAGTGTC CCGCTGAGGC CTGAGCGAGT GTCCAGCCAA GGGCTGAGTG TCCAGCACAC CTGCCGTCTT 21630  
 CACTTCCCCA CAGGCTGGCG CTCGGCTCCA CCCCAGGGCC AGCTTTTCTT CACCAGGAGC CCGGCTTCCA 21700  
 CTCGCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC GCCCTGCCCT CTTTGCCTT 21770  
 CCACCCCCAC CATCCAGGTG GAGACCCTGA GAAGGACCCT GGGAGCTCTG GGAATTTGGA GTGACCAAAG 21840  
 GTGTGCCCTG TACACAGGCG AGGACCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT GGGGGGAGGT 21910  
 GTGTGGGAG TAAATACTG AATATATGAG TTTTTCAGTT TTGAAAAAAA TCTCATGTTT GAATCTAAT 21980  
 GTGCACTGCA TAGACACCAC TGTATGCAAT TACAGAAGCC TGTGAGTGAA CGGGGTGGTG GTCAGTGGCG 22050  
 GCCCATGGCC TGGCTGTGCA TTTACGGAAG TCTATGAGTG AATGGGTTG TGGTCACTGC GGGCCCATGG 22120  
 CCTGGCTGGG CTTGGGAGGT TCTGATGCT GTGAGGCAGG AGGGGAAGGA GGGTAGGGGA TAGACAGTGG 22190  
 GAGCCCCCAC CTTGGAAGAC ATAACAGTAA GTCCAGGCCC GAAGGGCAGC AGGGATGCTG GGGGCCACGC 22260  
 TTGGGCGGCG GGGATGATGG AGGGCCTGGC CAGGGTGGCA GGGATGATGG GGGCCCCAGC TGGGTGGCA 22330  
 GGGGTGATGG GGGGGGCTGG TCTGGGTGGC GGGGAAGATG GGGGAAGCCTG GCTGGGCCCC CTCTCCCTT 22400  
 GCCTCCACCC TGCAGCCGTG GATCCGGATG TGCTTCCCTG GTGCACATCC TCTGGCCAT CAGTTTTCAT 22470  
 GGAGGTGGGG GGCAGGGGCA TGACACCATC CTGTATAAAA TCCAGGATTC CTCTCTCTGA ACGCCCCAAC 22540  
 TCAGGTTGAA AGTCACATTC CGCCTCTGGC CATCTCTTAA AGAGTAGACC AGGATCTCTGA TCTCTGAAGG 22610  
 GTGGGTAGGG TGGGGCAGTG GAGGGTGTGG ACACAGGAGG CTTAGGGTG GGGCTGGTGA TGCTCTCTCA 22680  
 TCCTCTTATC ATCTCCAGT CTCTCTCTCT ATCTCTTAT CATCTCCAG TCTCATCTGT CTCTCTCTTA 22750  
 TCTCCAGTC TCATCTGTCA TCCTCTTACC ATCTCCAGT CTCTCTCTT ATCTCTTAT CTCTAGTCT 22820  
 CATCCAGACT TACCTCCAG GCGGGGTGCC AGGCTCGCAG TGGAGCTGGA CATACGTCTT TCCTCAGGCA 22890  
 GAAGGAACCTG GAAGGATTGC AGAGAACAGG AGGGGCGGCT CAGAGGGAGC CAGTCTTGGG GTGAAGAAAC 22960  
 AGCCCTCCT CAGAAGTTGG CTTGGGCCAC ACGAAACCGA GGGCCCTCGG TGAGTGGCTC CAGAGCCTTC 23030  
 CAGCAGGTCC CTGGTGGGCG CTTATGGTAT GGGCGGGTCC TACTGAGTGC ACCTTGGACA GGGCTCTCGG 23100  
 TTTGAGTGCA GCGCGGAGCT GCGTGGTGTG GGGGTGGGGG CTTATGGCCA CTGGATATGG CGTCACTTAT 23170  
 TGCTGCTGCT TCAGAGAAATG TCTGAGTGAC CGAGCCTAAT GTGTATGGTG GGGCCAGTGC CACAGACTGT 23240  
 GTCGTAAATG CACTCTGGTG CTTGGAGCCC CCGTATAGGA GCTGTGAGGA AGGAGGGGCT CTTGGCAGCC 23310  
 GGCCTGGGGG CGCCTTTGCC CTGCAAACTG GAAGGGAGCG GCGCCGGGCG CCGTGGGCGG ACGACCTCAA 23380  
 GTGAGAGGTT GGACAGAACA GGGCGGGGAC TTCCAGGAG CAGAGGCCGC TGCTCAGGCA CACCTGGGTT 23450  
 TGAATCACAG ACCAACAGGT CAGGCCATTG TTCAGCTATC CATCTTCTAC AAAGCTCCAG ATTCTGTTT 23520  
 CTCCGGGTGT TTTTGTGTA AATTTTACTC AGGATTACTT ATATTTTGT CTAAGATATT AGACCCTTAA 23590  
 AAAAGGTATT TGCTTTGATA TGCTTAACT CACTAAGCAC CTACTTTATT TGCTGTTTT TATTTATTAT 23660  
 TATTATTATT ATTAGAGATG GTGTCTACTC TGTCACCCAG GTTGTAGTG CAGTGGCACA GTCATGGCTC 23730  
 GCTGTAGCCG CAAACCCCA GGCTCAAGTG ATCTCCGGC CTCAGCTTCC CAGAGTGTG GATTACAGG 23800  
 TGTGAGCCAC TGCCCTTGCC TGGCACTTTT AAAAACCACT ATGTAAGGTC AGGTCCAGTG GCTTCCACAC 23870  
 CTGTATCCC AGTAGTTTGG GAAGCCGAGG CAGAAGGATT GTCTGAGGCC AGGAGTTTGA GACCAGCATG 23940  
 GGTAACATAG GGAGACCCCA TCTCTACAAA AAATGCAAAA AGTTATCCGG GCGTGGGGTC CAGCATCTGT 24010  
 AGTCCAGCT GCTCGGGAGG CTGAGTGGGA GGATCGCTTG AGCCCGGGAG GTCATGGCTG CAGTGAGCTG 24080  
 TGATTGTACC ATCGCACTCC AGCCTGGGCA ACAGAGTGAG ACCCTGTCTC AAAAAAAAAA AAAAAAAAAA 24150  
 AAGGAGAAGG AGAAGAGAAG AAGAAGGAAG AAGGAAAGAG AAGAAGAAGG AAGAAGGAAG AAGAAGGAG 24220  
 AAGGAGGCC TTAGGTGCT AGGTAGACTG TCAATCTCA GAGCAAAATG AAAATAACAA AGTTTAAAG 24290  
 GGAAAGAAAA ACCCCAGCTC TTTGGACTTC CTTAGGCCCTG AACTTCATCT CAAGCAGCTT CCTTCCACAG 24360  
 ACAAGCGTGT ATGGAGCGAG TGAGTTCAAA CGAGAAAGGG AGGAGAAGCA GGAAGGGTG GAGGCTGTGG 24430  
 GTGACACCAG CCAGGACCCC TGAAGGGAG TGGTTGTTTT CTGCTCAG CCCCACGCTC CTGCCGGTCC 24500  
 TGCACCTGCT GTAACCGTCG ATGTTGGTGC CAGGTGCCCA CTTGGGAAGG ATGCTGTGCA GGGGCTTGC 24570  
 CAAACTTTGG TGGTTTTCAG AAGCCCCAGG CACTTGTGGC AGGCACAATT ACAGCCCTC CCCAAAGATG 24640  
 CCCACGTCTT TCTCTGGAA CCGTGAATG TGTACCCCGC AAGGCAGAGG CTGGTGAAAG CTGCAGGTGG 24710  
 AATCAGCGCT GCCAGTCAGC CGATCTTAAG GTCATCTGG ATTATCTGGT GGGCCTGATA TGGCCACAAG 24780  
 GGTCCCTAGA AGTGAGAGAG GGAGGCAGGG GAGAGTCAGA GAGGGGACGT GAGAAGGACC ACTGGCCACT 24850  
 GCTGGCTTTG AGATGGAGGA GGGGGTCCCC AGCCAAGGAA TGGGGGACG CGCTCCATGC TGGAAAAGCA 24920  
 AGCAATCCTC CCCGTCCTG AGGGCACACG CCGCTGCCCA CGCCTCGATT TCAGGCCAGT GGGACCTGTT 24990  
 TCAGCTTTCC GGCCTCCAGA GCTGTAAAGT GATGCGTTTG TGTTCAGCCA CTAAGCTGCA GTGATTCTG 25060  
 ACAGCAGCAA ATGGAATAGC AGTACAGGGA AATGAATACA GGGACAGTTC TCAGAGTGAC TCTCAGCCCA 25130  
 CCCCTGGG

## 60 Beispiel 5

Der Vergleich der oben beschriebenen genomischen hTC-Sequenz mit der Sequenz  
 der hTC-cDNA (Fig. 6; entsprechend SEQ ID NO 2) ermöglichte die Aufklärung der  
 Exon-Intron-Struktur des hTC-Gens. Die genomische Organisation des hTC-Gens ist  
 in Fig. 7 schematisch dargestellt. Die kodierende Region des hTC-Gens setzt sich aus

16 Exons zusammen, die in ihrer Größe zwischen 62 bp und 1354 bp variieren (s. Tabelle 1). Exon 1 enthält das Translationsstartcodon ATG. Das Translationsstopcodon TGA sowie der 3' untranslatierte Bereich liegen auf Exon 16 (Fig. 8). Ein mögliches Polyadenylierungssignal (AATAAA) wurde weder in Exon 16 noch in den 3195 bp der folgenden 3'-flankierenden Region gefunden. Basierend auf der Konsensussequenz

	5'-Exon			Intron				3'-Exon			
Prä-mRNA	A/C	A	G		G	T	A/G	A	...	N C	A G   G
Häufigk.(%)	70	60	80		100	100	95	70		80	100 100 60

wurden die Exon-Intron-Übergänge bestimmt und in Tabelle 1 aufgeführt. Mit Ausnahme der 5'-Splice-Stelle zwischen Exon 15 und Intron 15 stimmen alle Exon-Intron-Übergänge mit der publizierten (Shapiro und Senapathy, 1987) Splice-Konsensussequenz überein. Die Größe der Introns liegt zwischen 104 bp und 8616 bp. Da Intron 6 nur zum Teil isoliert wurde, kann die exakte Länge des hTC-Gens nicht bestimmt werden. Basierend auf der von Intron 6 erhaltenen Teilsequenz von ~4660 bp beträgt die minimale Größe des hTERT Gens 37 kb.

Die Introns 1-5 sowie der 5'-Bereich des Introns 6 sind in Contig 1 enthalten:

Intron 1: bp 11493-11596 (SEQ ID NO 4);

Intron 2: bp 12951-21566 (SEQ ID NO 5);

Intron 3: bp 21763-23851 (SEQ ID NO 6);

5 Intron 4: bp 24033-24719 (SEQ ID NO 7);

Intron 5: bp 24900-25393 (SEQ ID NO 8);

5'-Bereich von Intron 6: bp 25550-26414 (SEQ ID NO 9).

Der 3'-Bereich des Introns 6 sowie die Introns 7-15 sind in Contig 2 an folgenden

10 Positionen lokalisiert:

3'-Bereich von Intron 6: bp 1-3782 (SEQ ID NO 10);

Intron 7: bp 3879-4858 (SEQ ID NO 11);

Intron 8: bp 4945-7429 (SEQ ID NO 12);

Intron 9: bp 7544-9527 (SEQ ID NO 13);

15 Intron 10: bp 9600-11470 (SEQ ID NO 14);

Intron 11: bp 11660-15460 (SEQ ID NO 15);

Intron 12: bp 15588-16467 (SEQ ID NO 16);

Intron 13: bp 16530-19715 (SEQ ID NO 17);

Intron 14: 19841-20621 (SEQ ID NO 18);

20 Intron 15: 20760-21295 (SEQ ID NO 19).

Der 3'-nichttranskribierte Bereich befindet sich ebenfalls im Contig 2 an Position  
21960-25138 (SEQ ID NO 20).

25 Die genannten Introns haben im einzelnen folgende Sequenzen:

## Intron 1 (SEQ ID NO 4)

GTGGGCCTCCCCGGGGTCGGCGTCCGGCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGG  
CGACTCAGGGCGCTTCCCCCGCAG

5

## Intron 2 (SEQ ID NO 5)

GTGAGGAGGTGGTGGCCGTCGAGGGCCCCAGGCCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCC  
CTGGTCTCCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACGGTGATCTCTGCC  
TCTGCTCTCCCTCCTGTCCAGTTTGCATAAACTTACGAGGTTACCTTCACGTTTTGATGGACACGCGGTTTCCAGGCGC  
10 CGAGGCCAGAGCAGTGAACAGAGGAGGCTGGGCGCGGCAGTGGAGCCGGTTGCCGCAATGGGAGAGAAGTGTCTGGAAG  
CACAGACGCTCTGGCGAGGGTGCTGCAGGTTACCTATAATCCTCTTCGCAATTTCAAGGGTGGGAATGAGAGGTGGGA  
CGAGAACCCCTCTTCTGGGGTGGGAGGTAAAGGTTTTCAGGTGCACGTGGTCAGCCAATATGACGGTTTGTGTTTA  
AGATTTAATTGTGTGTTGACGGCCAGGTGCGGTGGCTCACGCCGTAATCCAGCACTTTGGGAAGCTGAGGCAGGTGGA  
TCACCTGAGGTGAGGAGTTTGAGACCAGCCTGACCAACATGGTGAAACCTATCTGTACTAAAAATACAAAAATTAGCTG  
15 GGCATGGTGGTGTGTGCCTGTAATCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCACTTGAACCCAGGAGGCGGAGGC  
TGCAGTGAGCTGAGATTGTGCCATTGTACTCCAGCCTGGGCGACAAGAGTGAACTCTGTCTTTAAAAAAGTGT  
CGTTGATTGTGCCAGGACAGGGTAGAGGAGGAGATAAGACTGTTCTCCAGCACAGATCCTGGTCCCATCTTTAGGTAT  
GAAGAGGGCCACATGGGAGCAGAGGACAGCAGATGGCTCCACCTGCTGAGGAAGGGACAGTGTGTTGTGGGTGTTCCAGGG  
ATGGTGTGCTGGGCCCTGCGGTGTCCTCCACCTGTTTTTCTGGATTGATGTTGAGGAACCTCCGCTCCAGCCCCCTTT  
20 TGGCTCCAGTGCTCCAGGCCCTACCGTGGCAGCTAGAAGAAGTCCCGATTTCACCCCTCCCCACAACTCCCAAGAC  
ATGTAAGACTTCCGGCCATGCAGACAAGGAGGGTGACCTTCTTGGGGCTCTTTTTTTCTTTTTTCTTTTATGGTGGC  
AAAAGTCATATAACATGAGATTGGCACTCCTAACACCGTTTTCTGTGTACAGTGCAAAATGCTAACTCGGCGGTGTTA  
CAGCAGGTTGCTTGAAATGCTGCGTCTTGGTGAAGTCCCTACCCATCGAACGGCAGCTGCCTCACACCTGCTGC  
GGCTCAGGTGGACCACGCCGAGTCAGATAAGCGTCATGCAACCCAGTTTGTCTTTTGTGCTCCAGCTTCCTTCGTTGAG  
25 GAGAGTTTGAGTTCTCTGATCAGGACTCTGCCTGTCTGCTGTTCTCTGACTTCAGATGAGGTACAATCTGCCCTGG  
CTTATGCAGGGAGTGAGGCGTGGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGGCGTTGCCCCAGGTGTCCCT  
GTCAGTGTAGGGTGAGTGAGGCGCGGCCCGGGTGTCCCTGTCCCGTGCAGCGTGATTGAGGTGTGGCCCCGGGTGT  
CCCTGTACGTGTAGGGTGAGTGAGGCGCCATCCCCGGGTGTCCCTGTACGTGTAGGGTGAGTGAGGCGTGGTCCCCGG  
GTGTCCCTGTCCCGTGCAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGGCGCGGTCC  
30 CCGGGTGTCCCTCTCAGGTGTAGGGTGAGTGAGGCGCGGCCCCAGGGTGTCCCTGTACGTGTAGGGTGAGTGAGGCACC  
GTCCCTGGGTGTCCCTCCAGGTATAGGGTGAGTGAGGCACTGTCCCCGGGTGTCCCTGTACGTGCAGGGTGAGTGAGG  
CGCGGCCCCGGGTGTCCCTCTCAGGTGCAGGGTGAGTGAGGCGCTGTCCCTGGGTGTCCCTGTCTGTTAGGGTGAGT  
GAGGCTCTGTCCCGAGGTGTCTTGGCGTTTGCTCACTTGAGCTTGCTCCTGAATGTTTGCTCTTTCTATAGCCACAGCT  
GCGCGGGTTGCCATTGCTGGGTAGATGGTGCAGGCGCAGTGCTGGTCCCCAAGCCTATCTTTCTGATGCTCGGCTCT  
35 TCTTGGTCACTCTCCGTTCATTTTGCTACGGGGACACGGGACTGCAGGCTCTCGCCTCCCGGTGCCAGGCACTGCAG  
CCACAGCTTCAGGTCCGCTTGCTCTGTTGGGCTGGCTTGCTCACCACGTGCCCGCCACATGCATGCTGCCAATACTCC  
TCTCCAGCTTGTCTCATGCCGAGGCTGACTCTGGGCTGCTGTCTGCTGCCACGTGTTGCTGGAGACATCCAGAA  
AGGGTTCTCTGTGCCCTGAAGGAAAGCAAGTCACCCAGCCCCCTCACTTGTCTGTTTTCTCCCAAGCTGCCCTCTGC  
TTGGCCCCCTTGGGTGGGTGGCAACGCTTGTCACCTTATTCTGGGCACCTGCCGCTCATGTGTTAGGCTGGGCTCTGCCT  
40 CCAGTCGCCCCCTCACATGGATTGACGTCCAGCCACAGGTTGGAGTGTCTCTGTCTGTCTCTGCTCTGAGACCCACGTG

GAGGGCCGGTGTCTCCGCCAGCCTTCGTCAGACTTCCCTCTTGGGTCTTAGTTTTGAATTTCACTGATTTACCTCTGACG  
TTTCTATCTCTCCATTGTATGCTTTTTCTTGGTTTATCTTTCATTCCCTTTCTAGCTTCTTAGTTTAGTCATGCCTTTC  
CCTCTAAGTGCTGCCTTACCTGCACCCGTGTTTTGATGTGAAGTAATCTCAACATCAGCCACTTTCAAGTGTTCTTAAA  
ATACCTTCAAAGTGTTAATACTTCTTTAAGTATTCTTATTCTGTGATTTTTTTCTTTGTGTCAGCGTGTGTTTTGACGTGA  
5 AATCATTTTGTATATCAGTGACTTTTAAGTATTCTTTAGCTTATTCTGTGATTTCTTTGAGCAGTGAGTTATTTGAACACT  
GTTTATGTTCAAGATATGTAGAGTATCAAGATACGTAGAGTATTTTAAGTTATCATTTTATTATTGATTTCTAACTCAGT  
TGTGTAGTGGTCTGTATAATACCAATTATTTGAAGTTTGGGAGCCTTGCTTTGTGATCTAGTGTGTGCATGGTTCCAG  
AACTGTCCATTGTAAATTTGACATCCTGTCAATAGTGGGCATGCATGTTCACTATATCCAGCTTATTAAGGTCCAGTGCA  
AAGCTTCTGTCTCCTTCTAGATGCATGAAATTCAGAAGGAGGCCATAGTCCCTCACCTGGGGGATGGGTCTGTTTATT  
10 TCTTCTCGTTTGGTAGCATTTATGTGAGGCATTGTTAGGTGCATGCACGTGGTAGAATTTTATCTTCTGATGAGTGAA  
TCTTTTGGAGACTTCTATGTCTCTAGTAATCTAGTAATCTTTTTTAAATTGCTCTTAGTACTGCCACACTGGGCTTCT  
TTTGATTAGTATTTTCTGCTGTGTCTGTTTTCTGCCCTTAATTTATATATATATATATATTTTTTTTTTTTTTGGAGACA  
GAGTCTTGGTCTGTGCGCCAGGGTGAGTGAGTGGTGTGATCACAGGTGAGTGAACTTTACCTTCTGGCCTGAGCCGT  
CCTCTCACCTCAGCCTCTGAGTAGCTGGAACGCAGACACGACCGCTACACCTGGCTAATTTTAAATTTTTTCTGGA  
15 GACAGGGTCTTGCTGTGTGCCCAGGCTGGTCTCAAACCTCTGGACTCAAGGGATCCATCTACCTCGGCTTCCCAAAGTG  
CTGAATTACAGGCATGAGCCACCATGTCTGGCCTAATTTTCAACACTTTTATATTCTTATAGTGTGGGTATGCTCTGTTA  
ACAGCATGTAGGTGAATTTCCAATCCAGTCTGACAGTCGTTGTTTAACTGGATAACCTGATTATTTTCAATTTTTTGTG  
ACTAGAGACCCGCTGGTGCACTCTGATTCTCCACTTGCTGTTGCATGCTCGTTCCCTGTTTCTCACCACCTCTTG  
GGTTGCCATGTGCGTTTCTGCGAGTGTGTGTTGATCCTCTCGTTGCTCCTGGTCACTGGGCATTTGCTTTTATTCT  
20 CTTTGCTTAGTGTTACCCCTGATCTTTTATTGTGCTGTTGCTTTTGTGTTATTGAGACAGTCTCACTCTGTCAACCA  
GGCTGGAGTGAATGGCACAATCTCGGCTCACTGCAACCTCTGCTCCTCGGTTCAAGCAGTTCTCATTCCTCAACCTCA  
TGAGTAGCTGGGATTACAGGCGCCACCACCGCTGGCTAATTTTGTATTTTAGTAGAGATAGGCTTTCACCATGT  
TGGCCAGGCTGGTCTCAAACCTCTGACCTCAAGTGATCTGCCGCTTGGCCTCCACAGTGCTGGGATTACAGGTGCAA  
GCCACCGTGCCCGCATACCTTGATCTTTTAAATGAAGTCTGAAACATTGCTACCTTGTCTGAGCAATAAGACCTT  
25 AGTGATTTTAGCTCTGGCCACCCCGAGCTGTGTGCTGTTTTCCCTGCTGACTTAGTCTATCTCAGGCATCTTGACA  
CCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTCTGTAGCTTTGCCCGCCCTGCTTTTCTCTC  
TTTGTTCCCGTCTGTCTTCTGTCTCAGGCCGCGCTCTGGGTCCCCTTCTTGTCTTTGCGTGGTCTTCTGTCTTG  
TTATTGCTGGTAAACCCAGCTTTACCTGTGCTGGCCTCCATGGCATCTAGCGACGTCGGGGACCTCTGCTTATGATGC  
ACAGATGAAGATGTGGAGACTCACGAGGAGGGCGGTCACTTGGCCCGTGAGTGTCTGGAGCACCAGTGGCCAGCGTTC  
30 CTTAGCCAGTGAGTGACAGCAACGTCCGCTCGGCTGGGTTGAGCCTGGAAAAACCCAGGCATGTGGGGTCTGGTGGCT  
CCGCGGTGTCGAGTTTGAATCGCGCAAACCTGCGGTGTGGCGCCAGCTCTGACGGTGTGCTGCGGGGGAGTGTCTG  
CTTCTCTCCTTCTGCTTGGGAACAGGACAAAGGATGAGGCTCCGAGCCGTTGTGCGCCACAGGAGCATGACGTGAGCC  
ATGTGGATAATTTTAAATTTCTAGGCTGGGCGCGGTGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCAAGCGGG  
TGGATCACGAGGTGAGGAGTCTGAGACCATCTGGCCAACATGATGAAACCCATCTGTACTAAAAACAAAAAATTAGC  
35 TGGGCGTGGTGGCGGTGCTGTAATCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGTTGGAA  
GTTGCAGTGAGCCGACATTGCACCACTGCACTCCAGCCTGGCAACACAGCGAGACTCTGTCTCAAAAAAAAAAAAAA  
AAAAAAAAAAATTTAGTAGCCACATTAAAAAGTAAAAAGAAAAGGTGAAATTAATGTAATAATAGATTTTACTGAA  
GCCAGCATGTCCACACCTCATATTTAGGGTGTATTGGTGGGAGCATCACTCACAGGACATTTGACATTTTTTGGAGC  
TTTGTCTGCGGATCCCGTGTGTAGTCCCGTGTGCGGCATCTCGGCCTGGACCTGTGGGCTTCCATGCGCATGGCT  
40 GTTGTACCAGATGGTGCAGGTCCGGGATGAGGTGCGCCAGGCCCTCAGTGAGCTGGATGTGCAGTGTCCGGATGGTGCAG  
TCTGGGATGAGGTGCGCCAGGCCCTGCTGTGAGCTGGATGTGTGGTGTCTGGATGGTGCAGGTGAGGGTGGGTCTCCAG

5 G C C C T C G G T G A G C T G G A G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G T G A G G T C G C C A G G C C C T G C T G T G A G C T G G A T G  
T G T G G T G T C T G G A T G G T G C A G G T C A G G G T G A G G T C T C C A G G C C C T C G G T A A G C T G G A G G T A T G G A G T C C G G A T G A T G C A  
G G T C C G G G T G A G G T C G C C A G G C C C T G C T G T G A G C T G G A T G T G T G T G T C T G G A T G G T G C A G G T C T G G G G T G A G G T C A C C  
A G G C C C T G C G G T G A G C T G G G T G T G C G G T G T C T G G A T G G T G C A G G T C T G G A G T G A G G T C G C C A G A C G G T G C C A G A C C A T G C  
10 G G T G A G C T G G A T A T G C G G T G T C C G G A T G G T G C A G G T C T G G G G T G A G G T T G C C A G G C C C T G C T G T G A G T T G G A T G T G G G G T  
G T C C G G A T G C T G C A G G T C C G G T G T G A G G T C A C C A G G C C C T G C T G T G A G C T G G A T G T G T G T G T C T G G A T G G T G C A G G T C T  
G G G G T G A A G G T C G C C A G G C C C T G C T T G T G A G C T G G A T G T G T G T G T C T G G A T G G T G C A G G T C T G G A T G A G G T C G C C A G  
G C C C T C G G T G A G C T G G A T G T G C A G T G T C C A G A T G G T G C A G G T C C G G G T G A G G T C G C C A G A C C C T G C G G T G A G C T G G A T G  
T G C G G T G T C T G G A T G G T G C A G G T C T G G A T G A G G T C G C C A G G C C C T C G G T G A G C T G G A T G T A T G G A G T C C G G A T G G T G C C  
15 G G T C C G G G T G A G G T C G C C A G A C C C T G C T G T G A G C T G G A T G T G C G G T G T C T G G A T G G T A C A G G T C T G G A T G A G G T C G C C  
A G A C C C T G C T G T G A G C T G G A T A T G C G G T G T C C G G A T G G T G C A G G T C A G G G T G A G G T C T C C A G G C C C T C G G T G A G C T G G A  
G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G T G A G G T C G C C A G G C C C T G C T G T G A A C T G G A T G T G C G G C G T C T G G A T G G T  
G C A G G T C T G G G G T G T G G T C G C C A G G C C C T C G G T G A G C T G G A G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G T G A G G T C G  
C C A G G C C C T G C T G T G A G C T G G A T G T G C G G C G T C T G G A T G G T G C A G G T C T G G G G T G T G G T C G C C A G G C C C T C G G T G A G C T G  
20 G A G G T A T G G A G T C C G G A T G A T G C A G G T C C G G G T G A G G T T G C C A G G C C C T G C T G T G A G C T G G A T G T G T G T A T C C G G A T G  
G T G C A G T C C G G G T G A G G T C G C C A G G C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G G T G C A G G T C T G G G G T G A G G T  
C A C C A G G C C C T G C G G T G A G C T G G T T G T G C G G T G T C C G G T T G C T G C A G G T C C G G G T G A G T T C G C C A G G C C C T C G G T G A G C  
T G G A T G T G C G G T G T C C C C G T G T C C G G A T G G T G C A G G T C C A G G G T G A G G T C G C T A G G C C C T T G G T G G G G T G G A T G T G C C G T  
G T C C G G A T G G T G C A G G T C T G G G G T G A G G T C G C C A G G C C T T T G G T G A G C T G G A T G T G C G G T G T C T G C A T G G T G C A G G T C T G  
25 G G G T G A G G T C G C C A G G C C C T T G G T G G G C T G G A T G T G T G G T G T C C G G A T G G T G C A G G T C C G G C G T G A G G T C G C C A G G C C C T  
G C T G T G A G C T G G A T G T G C G G T G T C T G G A T G G T G C A G G T C C G G G T G A G G T A G C C A A G G C C T T C G G T G A G C T G G A T G T G G G  
G T G T C C G G A T G G T G C A G G T C C G G G T G A G G T C G C C A G G C C C T G C G G T T A G C T G G A T A T G C G G T G T C C G G A T G G T G C A G G T  
C C G G G T G A G G T C A C C A G G C C C T G C G G T T A G C T G G A T G T G C G G T G T C T G G A T G G T G C A G G T C C G G G T G A G G T C G C C A G G  
C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G G T G C A G G T C C G G G T G A G G T C G C C A G G C C C T G C A G T G A G C T G G A T G  
30 T G C T G T A T C C G G A T G G T G C A G G T C T G G C G T G A G G T C G C C A G G C C C T G C G G T T A G C T G G A T A T G C G G T G T C G G A T G G T G C A  
G G T C C G G G T G A G G T C A C C A G G C C C T G C G G T T A G C T G G A T G T G C G G T G T C C G G A T G G T G C A G G T C T G G G G T G A G G T C G C C  
A G G C C C T G C T G T G A G C T G G A T G T G C T G T A T C C G G A T G G T G C A G G T C C G G G T G A G G T C G C C A G G C C C T G C G G T G A G C T G G  
A T G T G C T G T A T C C G G A T G G T G C A G G T C T G G C G T G A G G T C G C C A G G C C C T G C G G T G A G C T G G A T G T G C A G T G T A C G G A T G G  
T G C A G G T C C G G G T G A G G T C G C C A G G C C C T G C G G T G G G C T G T A T G T G T G T T G T C T G G A T G G T G C A G G T C C G G G T G A G T T  
35 C G C C A G G C C C T G C G G T G A G C T G G A T G T G T G G T G T C T G G A T G C T G C A G G T C C G G G T G A G T T C G C C A G G C C C T C G G T G A G C  
T G G A T A T G C G G T G T C C C C G T G T C C G A A T G G T G C A G G T C C A G G G T G A G G T C G C C A G G C C C T T G G T G G G G T G G A T G T G C C G T  
G T C C G G A T G G T G C A G G T C T G G G G T G A G G T C G C C A G G C C C T T G G T G A G C T G G A T G T G C G G T G T C C G G A T G G T G C A G G T C C G  
G G G T G A G G T C A C C A G G C C C T C G G T G A T C T G G A T G T G G C A T G T C C T T C T C G T T T A A G

## 35 Intron 3 (SEQ ID NO 6)

40 G T A C T G T A T C C C C A C G C C A G G C C T C T G C T T C T C G A A G T C C T G G A A C A C C A G C C C G G C C T C A G C A T G C G C C T G T C T C C A C T  
T G C C T G T G C T T C C C T G G C T G T G C A G C T C T G G G C T G G G A G C C A G G G G C C C C G T C A C A G G C C T G G T C C A A G T G G A T T C T G T G  
C A A G G C T C T G A C T G C C T G G A G C T C A C G T T C T C T T A C T T G T A A A A T C A G G A G T T T G T G C C A A G T G G T C T C T A G G G T T T G T A  
A A G C A G A A G G G A T T T A A A T T A G A T G G A A C A C T A C C A T A G C C T C C T G C C T T T C C C T G G G A T G T G G G T C T G A T T C T C T C  
T C T C T T T T T T T T T C T T T T T G A G A T G G A G T C T C A C T C T G T T G C C C A G G C T G G A G T G C A G T G G C A T A A T C T T G G C T C A C T

5 GCAACCTCCACCTCCTGGGTTTAAAGCGATTACCCAGCCTCAGCCTCCTAAGTAGCTGGGATTACAGGCACCTGCCACCAC  
GCCTGGCTAATTTTGTACTTTTAGGAGAGACGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAACTCATGACCTCAGG  
TGATCCACCCACCTTGGCCTCCCAAAGTGCTGGGTTTACAGGCTAAGCCACCGTGCCAGCCCCGATTCTCTTTAATT  
CATGCTGTTCTGTATGAATCTTCAATCTATTGGATTTAGGTCATGAGAGGATAAAATCCACCCACTTGGCGACTCACTG  
10 CAGGGAGCACCTGTGCAGGGAGCACCTGGGGATAGGAGAGTTCCACCATGAGCTAACTTCTAGGTGGCTGCATTGAATG  
GCTGTGAGATTTTGTCTGCAATGTTTGGCTGATGAGAGTGTGAGATTGTGACAGATTCAAGCTGGATTGTCATCAGTGAG  
GGACGGGAGCGCTGGTCTGGGAGATGCCAGCCTGGCTGAGCCCAGGCCATGGTATTAGCTTCTCCGTGTCCCGCCAGGC  
TGACTGTGGAGGGCTTTAGTCAGAAGATCAGGGCTTCCCGAGCTCCCTGCACACTCGAGTCCCTGGGGGGCCTTGTGAC  
ACCCCATGCCCCAAATCAGGATGTCTGCAGAGGGAGCTGGCAGCAGACCTCGTCAGAGGTAACACAGCCTCTGGGCTGGG  
15 GACCCCGACGTGGTGTGGGGCCATTTCTTGCATCTGGGGAGGGTCAGGGCTTTCCTGTGGGAACAAGTTAATACAC  
AATGCACCTTACTTAGACTTTACACGTATTTAATGGTGTGCGACCAACATGGTCATTTGACCAGTATTTTGAAAGAAT  
TTAATTGGGGTGACCGGAAGGAGCAGACAGACGTGGTGGTCCCAAGATGCTCCTTGTCACTACTGGGACTGTTGTTCTG  
CCTGGGGGGCCTTGGAGGCCCTCCTCCCTGGACAGGGTACCGTGCCCTTTTCTACTCTGCTGGGCCTGCGGCCCTGCGGTC  
AGGGCACAGCTCCCGAGCACCCGCGCCCCAGTGTCCACGGAGTGCCAGGCTGTGAGCCACAGATGCCAGGTCCAGGT  
20 GTGGCCGCTCCAGCCCCGTGCCCCATGGTGGTTTTGGGGGAAAAGGCCAAGGGCAGAGGTGTGAGGAGTGGTGGG  
CTCATGAGAGCTGATTCTGCTCCTTGGCTGAGTGCCCTGAGCAGCCTCTCCCGCCCTCTCCATCTGAAGGATGTGGCT  
CTTTCTACCTGGGGGTCTGCTGGGGCCAGCCTTGGGTACCCAGTGGCTGTACCAGAGGGACAGGCATCCTGTGTGG  
AGGGGCATGGGTTACGTGGCCCCAGATGCAGCCTGGGACCAGGCTCCCTGGTGTGATGGTGGGACAGTACCCCTGGGG  
GTTGACCGCCGACTGGGCGTCCCAGGGTTGACTATAGGACCAGGTGTCCAGGTGCCCTGCAAGTAGAGGGGCTCTCAG  
25 AGGCGTCTGGCTGGCATGGGTGGACGTGGCCCCGGGCATGGCCTTCAGCGTGTGCTGCCGTGGGTGCCCTGAGCCCTCAC  
TGAGTCGGTGGGGCTTGTGGCTTCCCGTAGCTTCCCCCTAGTCTGTTGTCTGGCTGAGCAAGCCTCCTGAGGGGCTCT  
CTATTGCAG

#### Intron 4 (SEQ ID NO 7)

25 GTGGCTGTGCTTTGGTTTAACTTCCTTTTAAACAGAAGTGCCTTTGAGCCCCACATTTGGTATCAGCTTAGATGAAGGG  
CCCCGAGGAGGGGCCACGGGACACAGCCAGGGCCATGGCAGCGGCCAACCCATTGTGCGCAGAGTGGCCGAGG  
TGCCGGTGCCCTCAGAAAAGCAGCGTGGGGGTGTAGGGGGAGCTCCTGGGGCAGGGACAGGCTCTGAGGACCACAAGAAG  
CAGCCGGGCCAGGGCCTGGATGCAGCACGGCCGAGGTCTGGATCCGTGTCTGTGTTGGTGGCAGCCTCCGTGCGCT  
TCCGCTTACGGGGCCCGGGACAGGCCAGACTGCCAGGAGCCACCGGGCTCTGAGGATCCTGGACCTTGGCCACGG  
30 CTCTGCACCCACCCCTGTGGTGGCTGGCTGGCTGGCTGACCCCGTCATCTGAGGAGAGTGTGGGGTGGGTGGACAGAG  
GTGTGGCATGAGGATCCCGTGTGCAACACACATGCGGCCAGGAACCCGTTTCAAACAGGGTCTGAGGAAGCTGGGAGGGG  
TTCTAGGTCCCGGTCTGGGTGGCTGGGGACACTGGGGAGGGGCTGCTTCTCCCTGGGTCCCTATGGTGGGGTGGGCAC  
TTGGCCGGATCCACTTTCCTGACTGTCTCCCATGCTGTCCCCGCCAG

#### 35 Intron 5 (SEQ ID NO 8)

GTGGGTGCCGGGGACCCCGTGAGCAGCCCTGCTGGACCTTGGGAGTGGCTGCCTGATTGGCACCTCATGTTGGGTGGAG  
GAGGTACTCCTGGTGGGCCGAGGGAGTGACGGTGACCTGTCACTGTTGAGGACACACCTGGCACCTAGGGTGGAGGC  
CTTCAGCCTTTCTGCAGCACATGGGGCGACTGTGCACCTGACTGCCCGGGCTCCTATTCCCAAGGAGGGTCCCACTG  
GATTCCAGTTTCCGTGAGAGAAGGAACCGCAACGGCTCAGCCACCAGGCCCGGTGCCTTGACCCCACTGCTGAGCCAG  
40 GGGTCTCCTGCTCCTGAGGCTCAGAGAGGGGACACAGCCCGCCTGCCCTTGGGGTCTGGAGTGGTGGGGTGGAGAGAG

5'-Bereich Intron 6 (SEO ID NO 9)

3'-Bereich Intron 6 (SEQ ID NO 10)

GTGTGGGATTGGTTTTCATGTGTGGGATAGGTGGGGATCTGTGGGATTGGTTTATAGTGGGGTGAACACAGAGAGTTCAAG  
GCGAGCTTTCTTCCTGTAGTGGGTCTGCAGGTGCTCCAACAGCTTTATTGAGGAGACCATATCTTCCTTTGAACTATGGT  
CGGGTTTATAGTAAGTCAGGGGTGTGGAGGCCTCCCCTGGGCTCCCTGTTCTGTTTCTTCCACTCTGGGGTCGTGTGGTG  
CCTGCTGTGGTGTGTGGCCGGTGGGCAGGGCTTCCAGGCCTCCTTGTGTTCAATTGGCCTGGATGTGGCCCTGGCTACGCT  
CCGTCCTTGGAATTCCCCTGCGAGTTGGAGGCTTTCTTTCTTTCTTTTTTCTTTCTTTTTTTTTTTTTTTTGATAACAGA  
GTCTCGCTCTTTTTTGCCACAGGCTGGAGTGGTTTGGCGTGACTTGGCTCACTGCAACCTGTGCTTCCTGAGTTCAAAGCA  
ATTCTCTTGCCTCAGCCTCCCAAGTAGCTGGAATTATAGGCGCCACCACCATGCTGACTAATTTTTGTAAATTTTAGTAG  
AGACGAGGTTTCTCCATGTGGCCAGGCTGGTCTCGAACTCTGACCTCAGGTGATCCTCCACCTCGGCCTCCCAAAGT  
GCTGGGATGACAGGTGTGAACCGCCGCCCGCCGAGACTCGCTTCTGCAGCTTCCGTGAGATCTGCAGCGATAGCTG  
CCTGCAGCCTTGGTGTGACAACCTCCGTTTTCTCTCCAGGTCTCGCTAGGGGTCTTTCATTTCATGACTCTCTTCA  
CAGAAGAGTTTCAGTGTGCTGATTTCCCGGCTGTTTCTGCGTAATTGGTGTCTGCTGTTTATCGATGGCCTCCTTCCA  
TTTCTTTTAGGCTTTGTTTATTGTGTTTTTCCGGCTCCTGAAGGAAAGTTTCGATTATGGATGTTTGAACCTTTCTTT  
TCTAAACAAGCATCTGAAGTTGCCGTTTTCCCTCTAAAGCAGGGATCCCAGGCCCCCTGGCTGTGGAGTGGCACCGGTCT  
GGGGCCTGTTAGGAACCCGGCGCACAGCGGAGGCTAGGTGGGGTGTGGGGAGCCAGCGTTCGCCCTGAGCCCCCGCCC  
TCTCAGATCAGCAGTGGCATGCGGTGCTCAGAGGCGCACACCCCTACTGAGAAGTGTGCGTGAGAGGGGTCTAGATTCT  
GTGCTCCTTATGGGAATCTAATGCCTGATGATCTGAGGTGGAACCGTTTGCTCCCAAACCATCCCTTCCCCACTGTG  
TCCTGTGGAAAAATCGTCTTCCAGGAAACAGTCCCTGGTACCACAATGGTTGGGGACCCTGTGCTAAAGACCTGCTTCA  
GCAGCCTCTCGTCAGTGTGATATATTGGCTTTTCTGTGTTGAGTCCAGAATAATTACGGATTTCTGTGATGCTTTCCGC  
CGACCTCAGACCCATGGGCTATTTGTGGGCGTGTGCTGCTCCTGGGTGGGAAGGGTGACGGCCCCATGTACCTTCTT  
GTTACTGCCTTCCAGGTGGTTCTCAGGGTTGAATCGTACTCGATGTGGTTTTAGCCCACGGCCCTGCCGCCAGCTCTG  
GGGGCTGGGGAACATGCTGAAGCACAGAGTCACCGTGCGCGTCTTTTGATGCCTCACAAGCTCGAGGCCTCCTGTGTCG  
TGTTAGTGTGTGTACGTGCTGCTCACATCCTGTCTTGGGACGACAGGGGCTTAGCAGGTCCCGTAGTAAATGACAAGC  
GTCCTGGGGGAGTCTGCAGAAATAGGAGGTGGGGGTGCGGTCTCTCTCCGCGTCTTCAGACTCTTCTCCTGCTGTGCTG

GTGGCTGCACCTGCATCCCTGCAATCCCTCCAGCACTGGGCTGGAGAGGCCCGGGAGCTCGAGTGCCACTTGTGCCACGT  
GACTGTGGATGGCAGTCGGTCACGGGGTCTGATGTGTGGTGACTGTGGATGGCGGTTGGTCACAGGGTCTGATGTGTG  
GTGACTGTGGATGGCGGTCGTGGGCTGATGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGG  
ATGGCGGTCGTGGGCTGATGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGGTGACTGTGGATGGCGGTCGTG  
5 GGGTCTGATGTGGTGACTGTGGATGGCAGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATG  
TGGTGACTGTGGATGGCAGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACT  
GTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGG  
CGGTCGTGGGCTGATGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGTGATCGGTCA  
CAGGGTCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGTGATCGGTACAG  
10 GGGTCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTTGGTCCCGGGG  
TCTGATGTGTGGTGACTGTGGATGGCGATCGGTACAGGGTCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCT  
GATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGGT  
GACTGTGGATGGCGGTCGTGGGCTGATGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGAT  
GGCGGTTGGTCCCGGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGGTGACTGTGGATGGCAG  
15 TCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGG  
TCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGT  
GGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGTGATCGGTACAGGGTCTGATGTGTGGT  
GACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGGTGACTGTGGAT  
GGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTAGGCTGATGTGTGGTGACTGTGGATGGCAGTCG  
20 GTACAGGGTCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGG  
GGTCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGATGTGTGGTGACTGTGGATGGCGGTCGTGGGCTGAT  
GTGGTGACTGTGGATGGTGATCGGTACAGGGTCTGATGTGTGGTAGCTGCAGGTGGAGTCCAGGTGTGTCTGTAGCT  
ACTTTGCGTCTCGGCCCCCGGCCCCGTTTCCAAACAGAAGCTTCCAGGCGCTCTCTGGGCTTCATCCCGCATCG  
GGCTTGGCCGAGGTCCACACGTCTGATCGGAAGAAACAAGTCCCAGCTCTGGCCGGGACAGGCCAATTTGTGGCTC  
25 ATGCCCTCTCTCTGCCGGCAG

## Intron 7 (SEQ ID NO 11)

GTCTGGCACTGCCCTGCAGGTTGGGCACGGACTCCAGCAGTGGTCTCTCCCTGGGCAATCACTGGGCTCATGACCG  
GACAGACTGTTGGCCCTGGGGGAGTGGGGGAATGAGCTGTGATGGGGCATGATGAGCTGTGTGCCTTGGCGAAATC  
30 TGAGCTGGGCCATGCCAGGCTGCGACAGCTGCTGCATTAGGCACCTGCTCACGTTTGACTGCGCGGCTCTCTCAGTT  
CCGAGTGCCCTTGTTCATGATTTGCTAAATGTCTTCTGCGCAGTTTGATCTTGAGGCCAAAGGAAAGGTGTCCCCCT  
CCTTTAGGAGGGCAGGCCATGTTTGGAGCGTGTCTGCCAGCTGGCCCTCAGTGCTGGGCTGAGGCCAAAGGAAACG  
TGTCCCCCTTCTTAGGAGGACGGGCGTGTGTGAGCCACGCCCCGCTGAGCGGCTCTCAGTGCTGGGCTGTGCCACGT  
GGCCCTGTGGCCCTTTGCAGATGTGGTCTGTCCACGTGGCCCTGTGGCTCTTTGCAGATGCCTGTAGCACTTGCTCGGC  
35 TCTAGGGGACAGTCGTGTCCACCGCATGAGGCTCAGAGACCTCTGGGCGAATTCCTTGGCTCCAGGGTGGGGGTGGAG  
GTGGCCTGGGCTGCTGGGACCCAGACCTGTGCCCGGACGCTGGGACGCACTCCTGGATCACATATGCCATCCGGGCCA  
CGGTGGGCTGTGTGGGTGTGAGCCAGCTGGACCCACAGGTGGCCAGAGGAGACGTTCTGTGTACACACTCTGCCTAA  
GCCCATGTGTGTCTGCAGAGACTCGGCCCGGCCAGCCACGATGGCCCTGCATTCCAGCCAGCCCGCACTTCATACA  
AACACTGACCCAAAAGGACGGAGGGTCTTGGCCACGTGTCTGCTGTCTCAGACCCACCGGCTCACTCCCATGTG  
40 TCTCCCGTCTGCTTTCGAG

## Intron 8 (SEQ ID NO 12)

GTGAGTCAGGTGGCCAGGTGCCATTGCCCTGCGGGTGGCTGGGCGGGCTGGCAGGGCTTCTGCTCACCTCTCTCCTGCCC  
CTTCCCCACTGNCCTTCTGCCCGGGGCCACCAGAGTCTCCTTTTCTGGCCCCCGCCCCCTCCGGCTCCTGGGCTGCAGGC  
5 TCCCCAGGCCCCGAAACATGGCTCGGCTTGGCGCAGCCGGAGCGGAGCAGGTGCCACACGAGGCTTGGAAATGGCAAGC  
GGGGTGTGGAGTTGCTCCTGCGTGGAGGACGAGGGGCGGGGGTGTGTCTGGGTCAAGTGTGCGCCGAGCGTTTGGAGCCT  
GCAGCTTGTGAGCTCCAAGTTACTACTGACGCTGGACACCCGGCTCTCACACGCTTGTATCTCTCTCTCCCGATACAAAA  
GGATTTTATCCGATTCTCATTCTGTCCCTGTGCTGTGACCCCGGAGGGCGGGGCTCTCTCTGTGACTAGATTT  
CCCATCTGAAAGTGCAGGGTTGACCGTGTAGTTTGTCTCCTCTCGGGGGGCTGTGGTGGCCATGGGGCAGGCGGCCTGG  
10 GAGAGCTGCCGTACACAGCCACTGGGTGAGCCACACTACGGTGGTAGAGCCACAGTGCCCTGGTGCCACATCACGTCCT  
CTGGATTTTAAAGTAAACACACACCTCCCGGAGGCATCTGCCTGCGACCTGTGTGTCCTGGGAGAGTGGTAGCAC  
GGAGGAAATTGCTGCACACTCAAGGTCATCAGCAAGGTCATCCGCAGTCAGGTGGAACGTGGAGGCTCTCTCTGGGATC  
GTCTCCAGCGGATAAAGGACTGTGCACAGCTTCGGAAGCTTTTATTTAAAAATATAACTATTAATTATTGCATTATAAGT  
AATCACTAATGGTATCAGCAATTATAATTTTATTAAAGTATAATTAGAAATATTAAGTAGTACACACGTTCTGAAAAA  
15 CACAAATTGCACATGGCAGCAGAGTGAATTTTGGCCGAGGGACACGTGTGCACATGTGTGAAGCGGCCCCAGGCCAC  
AGAATTCGCTGACAAAGTCACCTCCCCAGAGAAGCCACCACGGGCCTCCTTCGTGGTCGTGAATTTTATTAAGATGGATC  
AAGTCACGTACCGTCCACGTGTGGCAGGGCTTTGGGGAATGTGAGGTGATGACTGCGTCCCTCATGCCCTGACAGACAGGA  
GGTGAAGTGTGTCTGTCTCCTGTCCCTAGGACACGGACAGGCCCGAAGCTCTAGTCCCCATCGTGGTCCAGTTTGGCCTCTGA  
ATAAAAAAGTCTTCAAAACCTGTTGCCCCAAAACTAAGAACAGAGAGAGTTTCCCATCCCATGTGCTCACAGGGGCGTA  
20 TCTGCTTGCCTTGACTCGCTGGGCTGGCCGGACTCCTAGAGTTGGTGCGTGTGCTTCTGTGAAAAAGTGCAGTCCTCTT  
GCCCCATCACTGTGATATCTGCACCAGCAAGGAAAGCCTCTTTTCTTTCTTTCTTTTCTTTTGGAGACGGAAACGTCA  
CTGTTGTCTGCTGGGCTTGAGTGACGTGGCGGATCTCACTCACTGCAACCTCCGCCTCCCGGGTCCAGCATTTCTC  
CTGCCTCAGCCTCCCGAGCAGCTGAGATTACAGGCACCCACCCCTGCGCCTGGCTAATTTTGTATTTTAGTAGAGAG  
GGGTTTTTGCCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCACCTCGGCCTCCCAAAGTGTG  
25 GGATTACAGGTGTGAGCCATCACGCCAGCCGAAAGCCTCTTTTAAGGTGACCACCTATAGCGCTTCCCGAAAATAAC  
AGGTCTTGTCTTTTGCAGTAGGCTGCAAGCGTCTCTTAGCAACAGGAGTGGCGTCTGTGGGCTCTGGGGATGGCTGAGGG  
TCGCGTGGCAGCCATGCCTTCTGTGTGCACCTTTAGGTTCCACGGGGCTATTCTGCTCTCACTGTTTGTCTGAAAAACGCA  
CCCTTGGCATCCTTGTGTTGGAGAGTTTCTGCTTCTCGTGGTTCATGCTGAAACTAGGGGCAAGGTTGTATCCGTTGGCGC  
GCAGCGGCTACATGTAGGTCATGAGTCTTTCACCGTGGACAAATTCCTTGAAAAAAGAGAGTCCGGTTAAGCAT  
30 TCATTCCGGGTCAAGTGTCTGGTTCTGTGAATAAACTCTAAGATTTAAGAAACCTTAATGAAAGAAACCTTGATGATTC  
AGAGCAAGGATGTGGTCACACCTGTGGCTGGATCTGTTTCAGCCGCCAGTGCATGGTGAGAGTGGGAGCAGGGATTG  
TTTGTTCAGAGGTCTCATCTGGTATGTTTCTGAGGTGTTTGGCGGCTGAATGGTAGACGTGCGTTTGTGTGTATGAGGT  
TCTGTGTCTGTGTGGCTCGGTTTGTAGGTACGCATGTCCAGCACATGCCCTGCCCGTCTCTCACCTGTGTCTTCCCGC  
CCCAG

35

## Intron 9 (SEQ ID NO 13)

GTGAGGCCTCCTCTTCCCCAGGGGGGCTTGGGTGGGGTTGATTGCTTTTGATGCATTAGTGTTAATATCTGGTG  
TCTGGAGACCATGACTGCTCTGTCTTGAGGAACAGACAAGGTTGCAGCCCTTCTTGGTATGAAGCCGCAGGGAGGGG  
TTGCACAGCCTGAGGACTGCGGGCTCCACGAGGCTCTGTCCAGCGGCATGTCCAGAGGCCTCAGGGCTCAGCAGGGC  
40 GAGGGCCGCTGCCCTGCATGATGAGCATGTGAATTCAACACCGAGGAAGCACACCAGCTTCTGTACGTCACCCAGGTTT

CGTTAGGGTCCTTGGGGAGATGGGGCTGGTGCAGCCTGAGGCCCCACATCTCCAGCAGGCCCTCGACAGGTGGCCTGGA  
CTGGGCGCCTCTTCAGCCCATTTGCCATCCCACTTG CATGGGGTCTACACCCAAGGACGCACACACCTAAATATCGTGCC  
AACCTAATGTGGTTCAACTCAGCTGGCTTTTATTGACAGCAGTTACTTTTTTTTTTTAACTTTAAGTTCTAGGGTAC  
ATGTGCACGACGTGCAGGTTAGTTACATATGTATACATGTGCCATGTTGGTGTGCTGCACCCATTAACTCATCATTTACA  
5 TTAGGTATATCTCCTAATGTATCCCTCCCCACTCCCCCATCCCATGACAGGCCCTGGTGTGTGATGTTCCCCACCCTG  
TGCCAAGTGTCTCATTGTTTCACTTCCCACCTGTGAGTGAGAACATGTGGTGTGGTGTCTTTCTTCTGCAATAGTTT  
GCTCAGAGTGATGGTTTCCAGCTTCGTCCATGTCCCTACAAAGGACATGAACTCATCTTTTATGACTGCATAGTATT  
CCGTGGTGTATATGTGCCACATTTCTTAATCCAGTCTATCATCGATGGACATTGGGGTGGTTGCAAGTCTTTGCTACT  
GTGAATAGTGCCGCAATAAACATACGTGTGCATGTGTCTTTATAGCAGCATGATTATAATCCTTTGGGTATATACCCAG  
10 TAATGGGATGGCTGGGTCAAATGGTATTTCTAGTTCTAGATCCTTGAGGAATCACCACACTGTCTTCCACAATGGTTGAA  
CTAGTTTACACTCCCAACAGTGTAAGTGTCTGGTGTGAGAGGATGTGGACAGCAGTTATTTTTTATGAAAA  
TAGTATCACTGAACAAGCAGACAGTTAGTGAAGGATGCGTCAGGAAGCCTGCAGGCCACACAGCCATTCTCTCGAAGAC  
TCCGGGTTTTTCTGTGCATCTTTGAAACTCTAGCTCCAATTATAGCATGTACAGTGGATCAAGGTTCTTCTTATTAA  
GGTTCAAGTTCTAGATTGAAATAAGTTTATGTAACAGAAACAAAAATTTCTGTACACACAATTGCTCTGGGATTGGA  
15 GGAAAGTGTCTCGAGCTGGCGGCACACTGGTCAGCCCTCTGGGACAGGATACCTCTGGCCCATGGTCATGGGCGCTGG  
GCTTGGGCTGAGGGTCACACAGTGCACCATGCCAGCTTCTGTGGATAGGATCTGGGTCTCGGATCATGCTGAGGACC  
ACAGCTGCCATGCTGGTAAAGGGCACCACGTGGCTCAGAGGGGCGAGGTTCCAGCCCCAGCTTTCTTACCGTCTTCAG  
TTATTTTCCCTAAGAGTCTGAGAAGTGGGCGCGCCTGATGGCCTTCGTTCTGCTTTCAGCTGGCACAGAATTGCACAA  
GCTGATGGTAAACACTGAGTACTTATAATGAATGAGGAATTGCTGTAGCAGTTAACTGTAGAGAGCTCGTCTGTTGGAAA  
20 GAAATTTAAGTTTTTCAATTAACCGCTTTGGAGAATGTTACTTTATTTATGGCTGTGTAATTTGTTTGACATTCACTCCC  
TCGTAGACAGATACTACGTAAAAAGTGTAAGTTAACCTTGCTGTGATTTTCCCTTATTTTAG

#### Intron 10 (SEQ ID NO 14)

GTGAGGCCCGTGCCGTGTGTCTGTGGGACCTCCACAGCCTGTGGGCTTTGCAGTTGAGCCCCCGTGTCTGCCCCTGG  
25 CACCGCAGCGTTGTCTCTGCCAAGTCTCTCTCTGCGGGTGTGGATCCGCAAGAGCAGAGGCGCTTGGCCGTGCACC  
CAGGCCTGGGGCGCAGGGGCACCTTCGGGAGGGAGTGGGTACCGTGCAGGCCCTGGTCTGCAGAGACGCACCCAGGTT  
ACACACGTGGTGAGTGCAGGCGGTGACCTGGCTCCTGCTGCTCTTTGAAAGTCAAGAGTGGCGGCTCCTGGGGCCCCAG  
TGAGACCCCAGGAGCTGTGCACAGGGCTGCAGGGCCGAGGCGGCAGCCTCTCCCCAGGGTGACCTGAGCCTGCGGA  
GAGCAGGAGCTGCTGAGTGAGCTGGCCACAGCGTTTCGCTGCGGTACGTTCTGCGTGGGGTTGTTTGGGATCGGTGGG  
30 AGAATTTGGATTGCTGAGTGCTGCTGTCTTGAACCACGGAGATGGCTAGGAGTGGGTTTTCAGAGTTGATTTTGTGAAT  
CAAATAAAATCAGGCACAGGGGACCTGGCCTCAGCACAGGGGATTGTCCAATGTGGTCCCCCTCAAGGGCGCCCCACAG  
AGCCGGTGGGCTTGTTTTAAAGTGCATTTGACGAGGACGAGAAACCTTGAAAGCTGTAAAGGGAACCTCAGAAAATG  
TGGCCGCCAGGGGTGGTTTCAGGTGCTTTGCTGGGCTGTGTTTGTGAAAACCAATTTGGACCCGCCCTCCAAGTCCACCC  
TCCAGGTCCACCTCCAGGGCCGCCCTGGGCTGGGGGTATGCCTGGCGTTCTTGTGCCGCGAGCCCGAGCACAGCAGGC  
35 TGTGCACATTTAAATCCACTAAGATTCACTCGGGGGAGCCAGGTCCCAAGCAACTGAGGGCTCAGGAGTCTGAGGCT  
GCTGAGGGGACAGAGCAGACGGGAACGCTGCTTCTGTGTGGCAAGTTCTGAGGGTGTGCGCAGGAGGTGGCTCAGA  
GTGTATGTTGGGGTCCACCGGGGGCAGAACTCTGTCTCTGATGAGTCGGCAGCCATGTAAACAGGAAGGGGTGGCCACAG  
GGAGCTGGGAATGCACAGGGGAGCTGCGCAGCTGGCCGAGGTCCAGGGCCAGGCCACAGGAAGGGCAGGGGGACGCCC  
GGGGCCACAGCAGAGGCCGAGGAAGGGAAGGGGATGCCAGGCCAGAGCAGAGGCTACCGGGCACAGGGGGCTCCCTG  
40 AGCTGGGTGAGCGAGGCTCATGACTCGGCGAGGGAACCTCCTTGACGTGAAGCTGACGACTGGTGTGCCAGCTCACAG

CCCAGCCAGGTCCCGCGCTGAGCAGGAACTCAGAACCCTCCCCTTTGTCTAAAGCACAGCAGATGCCTTCAGGGCATCT  
AGGAGAAAACAGGCAAAGTCGTGAGAAACGTCTTAAAGAAGGTGGGATGGTGGCAATTTCTTGTCCAGATTTTAGTCT  
GCCCCGACCACAGATGAGTCTATAACGGGATTGTGGTGTGCCATGGGGACACATGAGATGGACCATCACAGAGGCCAC  
TGGGGCTGCACCTCCCATCTGAGTCTGGCTGTCCCGGGTCCAGGCCAGGTTCTTGCATGCTCACCTACCTGTCTCTGCC  
5 GGGAGACAGGGAAAGCACCCGAAGTCTGGAGCAGGGCTGGGTCCAGGCTCCTCAGAGCTCCTGCCAGGCCAGCACCTT  
GCTCCAAATCACCATTCTCTGGGGTTTTCCAAAGCATTTAAACAAGGGTGTCAAGTTACCTCCTGGGTGACGGCCCCGCA  
TCCTGGGGCTGACATTGCCCCCTCTGCCTTAG

Intron 11 (SEQ ID NO 15)

10 GTGAGCGCACCTGGCCGGAAGTGGAGCCTGTGCCCGGTGGGGCAGGTGTGCTGCAGGGCCGTTGCGTCCACCTCTGCT  
TCCGTGTGGGGCAGGCGAGTGCCTCCCAAGGGTCCAGAGGCCACAGGGTGCCCTCGTCCCATCTGGGGCTGAGCAGA  
AATGCATCTTTCTGTGGGAGTGAGGGTGCTCACACGGGAGCAGTTTTCTGTGCTATTTTGGTAAAAGGAAATGGTGAC  
CAGACCTGGGTGCACTGAGGTGTCTTCAGAAAGCAGTCTGGATCCGAACCAAGACGCCCGGGCCCTGCTGGGCGTGAGT  
CTCTCAAACCGAACACAGGGGCCCTGCTGGGCATGAGTCCCTCTGAACCCGAGACCTGGGGCCCTGCTGGGCGTGAGT  
15 CTCTCCGAACCCAGAGACTTCAGGGCCCTTTTGGGCGTGAGTCTCTCCGCTGTGAGCCCCACACTCCAAGGCTCATCCAC  
AGTCTACAGGATGCCATGAGTTCATGATCACGTGTGACCCATCAGGGGACAGGGCCATGGTGTGGGGGGGTCTCTACAA  
AATCTGGGGTCTTGTTCCTCCAGAGCCGAGAGCTCAAGGCCCCGTCTCAGGCTCAGACACAATGAATTGAAGATGGA  
CACAGATGCAGAAATCTGTGCTGTTCTTTTATGAATAAAAAGTATCAACATTCCAGGCAGGGCAAGGTGGCTCACACCT  
ATAATCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACTTGAGGCCAGGAGTTTGAAGCCAACTAACCAACATAGTG  
20 AAATTCATTTCTACTTAAAAAATACAAAATTAGCCTGGCCTGGTGGCACACGCTGTAGTCCCCGCTATGCGGGAGGC  
TGAGGCAGGAGAATCATTTGAACCCAGGAGGCAGAGGTTCAGTGAGCCGAGATCACACCACTGCCTCCAGCCTGGGCA  
ACAGAGTGAGACTTCATCTTAAAAAAGTATCAGCATTCCAAAACCATAGTGAGCAGGTGTTTTTTTATTC  
TGTCTTCGATAATATTTACTGGTGTGTGTAGAGGCCGAACTGGGGGTGCCTTCCTCTGAAAGGCACACCTTCATGG  
GAAGAGAAATAAGTGGTGAATGGTTGTTAAACCAGAGGTTTAAACTGGGGTCTCTGCTGTCTGAGTTAACAGTCCAGATC  
25 TGGACTTTGCCCTCTTCCAGAATGCTCCCTGGGGTTTGCTTCATGGGGAGCAGCAGGTGGACACCCTCGTGATGGGG  
GAGCAGCAGGTGCAGACGCCCTCATGATGGGGAGTGGCAGGTGCAGACACCCTTGTCATGGTGCCAGCATGTCCCTG  
TTGCAGCTCCCTCCCCACAAGGATGCCGGTCTCCTGTGTCTCCCAAGTCCCTGCTTCCCTCTCACAGCCTTACCTGGTC  
CTGGCCTCCACTGGCTTTGTCTGCATGATTTCCACATTTCTGGGCTCCCAGCACCTCTTCGCCTCTCCCAGGCACCTCT  
GCAGTGCTGGCCATACCAGTCAGCTGTGAAGTGTCCACTGCTTATTTTGCTCCCCATGAAATGTATTTTTTAGGACAGGC  
30 ACCCCTGGTTCAGCCTCTGGCACAGCATCAGTGAATGTTATTGAAGGACAAAGGACAGACAAACAAATCAGGAAAATGG  
GTTCTCTTAAACACATTGCAAAGCCACAGAGGCTAGTGAGGATGGGTGGGCATCAGGTATCAGATGTGGGTCCAATG  
CCAGAATATTCTGTGCTCCCAAAGGCCACTTGGTCAGAGTGTGTGCTTGAGAGGTGGCTCTAAAAGCTCAGCAGTGGAG  
GCAGTGGTTCGCCATACTCAGGGTGAAGTACATCCTCTGTGTCTGAAGTATACAGCAGAGGCTTGAAGGCATCTGGGA  
GAAGAAAACAGGCAAAATGATTAAGAAAAGTAAAAAGGAAAGTGGTAAGATGGGAATTTCTTGTCCAGATTTTAGTC  
35 TCCCAAACCACAGCTCAGATGGTAGAATGTGGTCAGAACTGATGGACAGAACAAATAGAACAAAACGGAAGCCCTATCTCT  
CAGAAAACGTGTGTTAATGTGGTATGTGGCACAGCTGATGGAAAAGAGAGTGTGTGTGTAATTTTTTTCTGAGAAAAC  
GACTGGAAGCAAATAAGTTGTGTCTTTACAGCATATACCAGAGCAGATTCTAGGTAGAAGAGGAGACATGCAAAACAAC  
ACCAGCAACAGAAATAAAACAAAAGACTCAAAGGGAAGGGAGGTGAACGTTCCCTGGTTTGGTGTGGGGAAGGACACAC  
AGGGAGGCGGATGAAACAGTGAGGCAACGGGCATTGCTTTCACTGCAGAGAACTCAGCTTGCTGAGCCACAGTGAAA  
40 ATGGCCATTCCCTGGAGCGTTTGTGCACGTGATTTATTTAAGGCGCCCTGTGAGGTCTGCACATTCATCTCTCACTTT

5 GTTCTCCTAACCACTGAGAGGTAGAGGAGGAAAGGCTCCAGGGGAGCAGCCGCCCTTGGTCAACCAGCTGGCAAAGGGC  
ATGCATGATTGCAGCCTGGCCTCCTGCTCCGGGGCCCTTGCTCTGCCCCAGGACCCACACAAGTCAGACCCATAGGCTC  
AGGGTGAGCCGGAGCCCAAGGTCGTGTTGGGATGGCTGTGAAAGAAGAAATGGACGTCTGATGCACACTTGGGAAGGTC  
CTACCAGCAGCGTCAAAGAAATGCATGTGAACTGACAGCGAGACCCATCCCTCAAAGAAACGCACGTGAAACTGATGGC  
10 GAGACCTGTCCCCATCCCTCATGTGGCTCCTTTCTGGGCTTGCCAAGAGCCAGCATCAGGTTGAGGCAAGCTGGAAAG  
ACTTTTCTGGAAAGCAGCTTGTTCATGGAAGTCCTCACAATGTCTGTGCTTCCAGTAATTCACCTTCTGAAGTGA  
CCAGACATTATCACGGGTCTTATTTACCATTTCAGTGTTCCAGGCAGGGGACTTGCCACAGCAAGTCACGAACCTGCC  
CAAATACAGGGCTAAGGAGATATTATGCATCACAAAACCTTGCTCTGCCATTAAACATTTTCAAAGAATTTTGAAGAAT  
GTTTAAATGGCACAAAACGTTTATTTCAATGTAGCAGTGTTCAAAGCTGGATGTAAAAGAACACACCCAGGAGCCTGCCG  
15 TGAATGTATGTGTGTTTCTTTGACATGGACATACATGGGCAGTGAGTGGTGGTGAGGCCCTGGAGGACATCGGTGG  
GATGCCTCCATCCTGCCCTCTGGAGACACCATGTGTGCCAGTGCACTCACTGGAGCCCTGTTAGCTGGTGCCACCTG  
GCTCTTCCATCCCTGAGATTCAAACACAGTGAGATTCCCCACGCCCAACTCAGTGTTCTCCACAAAAAACCTGAGTCAC  
ACCTGTGTTCACTCGAGGGACGCCCGGGAGCCAGGGTCCACAGTTTATTATGTGTTTTTGGCTGAGTTATGTGCAGATC  
TCAATCAGGGCAGATGATGAGTGACAAAACAGCGCCGTGCGAGGTTTGGATACACTCAACATCACTAGCCAGGTCTGGTG  
20 GAGTTTGGTCATGCAGAGTCTGGATGGCATGTAGCATTGGAGTCCATGGAGTGAGCACCAGCCCTCGGGCTGCAGC  
GCATGCCCGAGGCAGGACAAGGAAGCGGGAGGAAGGCAGGAGGCTCTTTGGAGCAAGCTTTCAGGAGGGGGCTGGGTGT  
GGGGCAGGCACCTGTGTCTGACATTCCCCCTGTGTCTCAG

#### Intron 12 (SEQ ID NO 16)

20 GTGAGCAGGCTGATGGTCAGCACAGAGTTCAGAGTTCAGSAGGTGTGTGCGCAAGTATGTGTGTGTGTGTGCGCGCT  
GCCTGCAAGGCTGATGGTGAAGTGGCTGCACGTAAGAGTGACATGTACGCATATACAGTGAGCAGATACATGTGTGCAT  
GTGTGTACATGAAGGCATGGCAGTGTGTGCACAGGTGTGCAAGGGCACAAGTGTGTGCATGCGAATGCACACCTGACA  
TGATGTGTGTTTCGTGCACAGTCGTGTGGGCATTACGTGAGGTGCATGCGTGTGGGTGTGCAGTGTGAGTAGCATGTGT  
GCACATAACATGTATTGAGGGGTCTCGTGTTCACCCCGCTAGGTCTCAGCACCAGTGCCACTCCTTACAGGATGAGAC  
25 GGGGTCCCAGGCCTTGGTGGGTGAGGCTCTGAAGCTGCAGCCCTGAGGGCATGTCCCATCTGGGCATCCGCGTCCACT  
CCCTCTCCTGTGGGCTTCTGTGTCCACTCCCCCTCTCCTGTGGGCATTACATCCACTCCACTCCCTCTCTCCTGTGGGC  
ATCCGCGTCCACTCCCCCTCTCTGTGGGCATCTGCGTCCACTCCCTCTCTGTGGGCATTGCGTCCACTCCCTCTCCT  
GGTTCCCTCTCTGTCTTGGCCGAGCCTCGGGGGCAGGCAGATGACACAGAGTCTTGACTCGCCAGGGTGGTTGCGCAGCTG  
CCGGGTGAGGGCCAGGCCGATTTCACTGGGAAGAGGGATAGTTTCTGTCAAATGTTCTCTTTCTTGTTCATCTGA  
30 ATGGATGATAAAGCAAAAAGTAAAACTTAAATCCAGAGAGGTTTCTACCGTTTCTCACTCTTTCTTGGCGACTCTAG

#### Intron 13 (SEQ ID NO 17)

35 GTGAGCCGCCACCAAGGGGTGAGGCCAGCCTCCAGGGACCCTCCGCGCTCTGCTCACCTCTGACCCGGGGCTTCACTT  
TGGAACCTCTGGGTTTTAGGGCAAGGAATGTCTTACGTTTTTCACTGGTGCTGCTGCCTGTGCACAGTTCTGTTGCGGTG  
GCTCTGTGCAAAGCACCTGTTCTCCATCTCTGGGTAGTGGTAGGAGCCGGTGTGGCCCCAGGTGTCCCACTGTGCTGT  
GCACTGGCCGTGGGACGTATGGAGGCCATCCAGGGCAGCAGGGCATGGGGTAAAGAGATGTTTATGGGGAGTCTTAG  
CAGAGGAGGCTGGGAAGGTGTCTGAACAGTAGATGGGAGATCAGATGCCCGAGGATTGCGGTCTCAGCAAAGAGGGCC  
GAGGTGGGTGCAGGTGAGGGTCGCTGGCCCCACCCCGGAAGGTGCAGCAGAGCTGTGGTCCCCACACAGCCCGGCCA  
GCACCTGTGCTCTGGGCATGGCTGTGCTCCTGGAACGTTCCCTGTCTGGCTGGTCAGGGGGTCCCCCTGCCAAGAATCG  
40 ACAACTTTATCACAGAGGAAGGGCCAATCTGTGGAGGCCACAGGGCCAGCTTCTGCCTGGAGTCAGGGCAGGTGGTGGC

ACAAGCCTCGGGGCTGTACCAAAGGGCAGTCGGGCACCACAGGCCCGGGCCTCCACCTCAACAGGCCTCCCGAGCCACTG  
GGAGCTGAATGCCAGGAGGCCGAAGCCCTCGCCCCATGAGGGCTGAGAAGGAGTGTGAGCATTTGTGTTACCCAGGGCCG  
AGGCTGCGCGAATTACCGTGACACTTGATGTGAAATGAGGTCGTCGTCTATCGTGAAACCCAGCAAGGGCTCACGGGA  
GAGTTTTCCATTACAAGGTCGTACCATGAAAATGGTTTTTAACCCGAGTGTCTGCGCCTTCATGCTCTGGCAGGGAGGGC  
5 AGAGCCACAGCTGCATGTTACCGCCTTTGCACCACTCCAGAGGCTTGGGACCAGGCTGTCTCAGTTCAGGGGTGCGTCC  
GGCTCAGACCGCCCTCCTCTCTGCTTCTCTCTCTGCTCTCAAATCTTCCCTCGTTTGCATCTCCCTGACGCGTGCTGGG  
CCCTCGTGCAAGCTGCTTGACTCCTTTCCGGAACCCCTTGGGGTGTGCTGGATACAGGTGCCACTGAGGACTGGAGGTGT  
CTGACACTGTGGTTGACCCAGGGTCCAGCTGGCGTGCTGGGGCCTCCTTGGGCCATGATGAGGTGAGAGGAGTTTTCC  
CAGGTGAAAACCTCTGGGAACTCCAGGGCCATGTGACCTGCCACCTGCTCCTCCCATATTAGCTCAGTCTTGTCTCTC  
10 ATTTCCCCACCAGGCTCTCTAGCTCCGAGGAGCTCCCGTAGAGGGCTTGGGCTCAGGGCAGGGCGGCTGAGTTTCCCCAC  
CCATGTGGGGACCCCTTGGGTAGTCGCTTGATTGGGTAGCCCTGAGGAGGCCGAGATGCGATGGGCCACGGGCCGTTTCCA  
AACACAGAGTCAGGCACGTGGAAGGCCAGGAATCCCTTCCCTCGAGGCAGGAGTGGGAGAACGGAGAGCTGGGCCCCG  
ATTTACGGCAGCCAGGCTGCAGTGGGCGAGGCTGTGGTGGTCCACGTGGCGCTGGGGCGGGGTCTGATTCAAATCCGC  
TGGGGCTCGGCCTTCTGCGCCGTGCTGGCCGCGCTCCACAGGGCTTGGGGTGGACGCCCCGACCTCTAGCAGGTGGC  
15 TATTTCTCCCTTTGGAAGAGAGCCCTCACCCATGCTAGGTGTTTCCCTCCTGGGTGAGGAGCGTGGCCGTGTGGCAACC  
CCGGGACCTTAGGCTTATTTATTTGTTTAAAAACATTCTGGGCTGGCTTCCGTTGTTGCTAAATGGGAAAAGACATCC  
CACCTCAGCAGAGTTACTGAGAGGCTGAAACGGGGTGTGGCTTACTGGTGTGATCTCAGGTCAATCCAGAAGTGGCT  
CAGGAAGTCAGTGAGACCAGGTACATGGGGGCTCAGGCAGTGGGTGAGATGAGGTACACGGGGGGCTCAGGCAGTGGGT  
GAGGCCAGGTACATGGGGGGCTCAGGCACTGGGTGAGATGAGGTACACGGGGGGCTCAGGCAGAGGGTACAGCCAGGTAC  
20 ACGGGGGCTCTGATCACACGCACATATGAGCACATGTGCACATGTGCTGTTTCATGGTAGCCAGGTCTGTGCACACCTGC  
CCCAAAGTCCCAGGAAGCTGAGAGGCCAAAGATGGAGGCTGACAGGGCTGGCGCGGTGGCTCACACCTGTAGTCCCAGCA  
CTTTGGGAGGCCGAGGCGAGAGGATCCCTTGAGCCAGGAGTTAAGACCAGCCTGAGCAACATAGTAGAACCCCATCTC  
TATGAAAAATAAAAAACAAAATTAGCTGAACATGGTGGTGTGCGCCTGTAGTTCCAATACTTGGGAGGCTGAAGTGGGAG  
GATCACTTGAGCCAGGAGGTGGAAGCTGCAGTGAGCTGAGATTGCACCACTGTACTGCAGCCTGGGTGACAGAGTGAGA  
25 GCCCATCTCAACAACAACAAAGAAGACTGACAAATGCAGTTTCTTGAAAGAAACATTTAGTAGGAACTTAACCTACACA  
CAGAAGCCAAGTCGGTGTCTCGGTGTGAGTGAGATGAGATGAGGTCTCACACCATCACCCAGACCCAGGGTTTATG  
CACCAAGGGGCGGGTGGCTCAGAAGGGATGCGCAGGACGTTGATATACGATGACATCAAGTTGTCTGACGAAGGGCAG  
GATTCATGATAAGTACCTGCTGGTACACAAGGAACAATGGATAAACTGGAAACCTTAGAGGCCCTCCCGAACAGGGGCT  
AATCAGAAGCCAGCATGGGGGGCTGGCATCCAGGATGGAGCTGCTTCAGCCTCCACATGCGTGTTCATACAGATGGTGCA  
30 CAGAAACGCAGTGATCTGTGCACACAGACACGCAGCTACTCGCACACACAAGCACACACAGACATGCATGCATGC  
ATCCGTGTGTGTGCACCTGTGCCATGAGGAAACCCATGCATGTGCATTATGCACGCACACAGGCACCGGTGGGCCCCAT  
GCCACACCCACGAGCACCGTCTGATTAGGAGGCTTTCCTCTGACGCTGTCCGCCATCCTCTCAG

#### Intron 14 (SEQ ID NO 18)

35 GTATGTGCAGGTGCCTGGCCTCAGTGGCAGCAGTGCTGCTGCTGTTAGTGTGTGAGGAGACTGAGTGAATCTGGG  
CTTAGGAAGTTCTTACCCCTTTTCGCATCAGGAAGTGGTTTAACCCAAACCACTGTGAGGCTCGTCTGCCCCCCTCTCGT  
GGGGTGAGCAGAGCACCTGATGGAAGGGACAGGAGCTGTCTGGGAGCTGCCATCCTTCCACCTTGCTCTGCTGGGGAA  
GCGCTGGGGGGCCTGGTCTCTCCTGTTTCCCCATGGTGGGATTGGGGGGCCTGGCCTCTCCTGTTTGGCCTGTGGTGG  
GATTGGGCTGTCTCCCGTCCATGGCACTTAGGGCCCTTGTGCAAACCCAGGCCAAGGGCTTAGGAGGAGGCCAGGCCAG  
40 GCTACCCACCCCTCTCAGGAGCAGAGGCCGCTATCACACGACAGAGCCCGCGCGCTCTCTGCTTCCAGTCACCG

TCCTCTGCCCCCTGGACACTTTGTCCAGCATCAGGGAGGTTTCTGATCCGTCTGAAATTCAAGCCATGTGGAACCTGCGGT  
CCTGAGCTTAACAGCTTCTACTTTCTGTTCTTCTGTGTTGTGGAAATTCACCTGGAGAAGCCGAAGAAAACATTTCGT  
TCGTGACTCCTGCGGTGCTTGGGTGGGACAGCCAGAGATGGAGCCACCCCGCAGACCGTCGGGTGTGGGCAGCTTTCCG  
GTGTCTCCTGGAGGGGAGCTGGGCTGGGCTGTGACTCCTCAGCCTCTGTTTTCCCCCAG

5

**Intron 15 (WEQ ID NO 19)**

GCAAGTGTGGGTGGAGGCCAGTGGGGCCCCACCTGCCCCAGGGGTCTATCCTTGAACGCCCTGTGTGGGGCGAGCAGCCTC  
AGATGCTGCTGAAGTGAGACGCCCCCGGGCCTGACCCTGGGGGCTGGAGCCACGCTGGCAGCCCTATGTGATTAAACG  
CTGGTGTCCCCAGGCCAGGAGCCTGGCAGGTTCCCAACTTCTGAACCCCTGCTTCCCATCTCAGGGGCGATGGCTCC  
10 CCACGCTTGGGAGCCTTCTGACCCCTGACCTGTGTCTCTCACAGCCTCTTCCCTGGCTGTGCCCTGAGCTCCTGGGGT  
CCTGAGCAAGTTCTCTCCCGCCCCGCGCTCCAGCGTCACTGGGCTGCTGTCTGCTCGCCCCGGTGGAGGGGTGTCTG  
TCCCTTCACTGAGGTTCCACAGCCAGGGCCAGAGGTGAGGCCCTGCTGCCCCGCCACCCACAGCTCTAGGAGGG  
TTGGAGGATGCCACCTCTGGCCTCTTCTGGAACGGAGTCTGATTTTGGCCCCGAG

15 **3'-untranskribierter Bereich (SEQ ID NO 20)**

ATCTCATGTTTGAATCCTAATGTGCACTGCATAGACACCACTGTATGCAATTACAGAAGCCTGTGAGTGAACGGGGTGGT  
GGTCAGTGGGGCCCATGGCCTGGCTGTGATTACGGAAGTCTATGAGTGAATGGGGTTGTGGTCAGTGGGGCCCATG  
GCCTGGCTGGGCCTGGGAGGTTTCTGATGCTGTGAGGCAGGAGGGAAGGAGGGTAGGGGATAGACAGTGGGAGCCCCA  
CCCTGGAAGACATAACAGTAAGTCCAGGCCGAAGGGCAGCAGGGATGCTGGGGGCCAGCTTGGGCGCGGGGATGATG  
20 GAGGGCCTGGCCAGGGTGGCAGGGATGATGGGGGCCAGCTGGGGTGGCAGGGGTATGGGGGGGGTGGTCTGGGTGG  
CGGGGAAGATGGGAAGCCTGGCTGGGCCCCCTCTCCCCCTGCTCCACCTGCAGCCGTGGATCCGGATGTGCTTCCCT  
GGTGACATCCTCTGGGCCATCAGCTTTATGAGGTGGGGGGCAGGGGCATGACACCATCCTGTATAAAATCCAGGATT  
CCTCCTCTGAACGCCCAACTCAGGTTGAAAGTCACATTCCGCTCTGGCCATTCTCTTAAGAGTAGACCAGGATTCTG  
ATCTCTGAAGGGTGGGTAGGGTGGGGCAGTGGAGGGTGTGGACACAGGAGGCTTCAGGGTGGGGCTGGTGATGCTCTC  
25 ATCTCTTATCATCTCCAGTCTCATCTCTATCCTCTTATCATCTCCAGTCTCATCTGTCTTCTCTTATCTCCAGT  
CTCATCTGTCTCTTACCATCTCCAGTCTCATCTCTTATCCTCTTATCTCCTAGTCTCATCCAGACTTACCTCCCA  
GGGCGGGTGCCAGGCTCGCAGTGGAGCTGGACATACGTCCTTCTCAGGCAGAAGGAAGTGAAGGATTGCAGAGAACAG  
GAGGGGCGGCTCAGAGGGACGAGTCTTGGGGTGAAGAAACAGCCCCCTCCTCAGAAGTTGGCTTGGGCCACAGAAACCG  
AGGGCCCTGCGTGAGTGGCTCCAGAGCCTTCCAGCAGGTCCCTGGTGGGGCCTTATGGTATGGCCGGGTCTACTGAGTG  
30 CACCTTGGACAGGGCTTCTGGTTTGTGAGTGAGCCCGACGTGCTGGTGTGCGGGTGGGGCTTATGGCCACTGGATATG  
GCGTCATTTATTGTGCTGCTTCAGAGAATGTCTGAGTGACCGAGCCTAATGTGTATGGTGGGCCCAAGTCCACAGACTG  
TGTCGTAATGCACTCTGGTGCTGGAGCCCCGTATAGGAGCTGTGAGGAAGGAGGGGCTCTTGGCAGCGGCCTGGGG  
GCGCCTTTGCCCTGCAAACTGGAAGGGAGCGGCCCCGGGCGCCGTGGGCGGACGACCTCAAGTGAGAGGTTGGACAGAAC  
AGGGCGGGGACTTCCAGGAGCAGAGGCCGCTGCTCAGGCACACCTGGGTTTGAATCACAGACCAACAGGTGAGGCCATT  
35 GTTCAGCTATCCATCTTCTACAAAGCTCCAGATTCTGTTTTCTCCGGGTGTTTTTTGTTGAAATTTTACTCAGGATTACT  
TATATTTTTTGCTAAAGTATTAGACCCTTAAAAAGGTATTTGCTTTGATATGGCTTAACCTACTAAGCACCTACTTTAT  
TTGTCTGTTTTATTATTATTATTATTATTATTAGAGATGGTGTCTACTCTGTCAACCCAGGTGTTAGTGACGTGGCAC  
AGTCATGGCTCGCTGTAGCCGCAACCCCCAGGCTCAAGTGATCTCCGGCCTCAGCTTCCAGAGTGTGGGATTACAG  
GTGTGAGCCACTGCCCCCTGGCTGGCACTTTAAAAAACCCTATGTAAGGTGAGGTCCAGTGGCTTCCACACCTGTCTATCC  
40 CAGTAGTTTGGGAAGCCGAGGCAGAAGGATTGTCTGAGGCCAGGAGTTTGAGACCAGCATGGGTAACATAGGGAGACCCC

ATCTCTACAAAAATGCAAAAAGTTATCCGGGCGTGGGGTCCAGCATCTGTAGTCCCAGCTGCTCGGGAGGCTGAGTGGG  
AGGATCGCTTGAGCCCGGGAGGTCATGGCTGCAGTGAGCTGTGATTGTACCATCGCACTCCAGCCTGGGCAACAGAGTGA  
GACCCTGTCTCAAAAAAAAAAAAAAAAAAAGAAGGAGAAGGAGAAGAAGAAGGAAGAAAGAGAAGAAGAAG  
GAAGAAGGAAGAAAGAAGGAGAAGGAGGCGCTGCTAGGTGCTAGGTAGACTGTCAAATCTCAGAGCAAAATGAAAATAACA  
5 AAGTTTTAAAGGGAAGAAAAACCCAGCTCTTTGGACTTCCTTAGGCCTGAACTTCATCTCAAGCAGCTTCCTTCCACA  
GACAAGCGTGTATGGAGCGAGTGAGTTCAAAGCAGAAAGGGAGGAGAAGCAGGCAAGGGTGGAGGCTGTGGGTGACACCA  
GCCAGGACCCCTGAAAGGGAGTGGTTGTTTTCTGCCTCAGCCCCACGCTCCTGCCGGTCTGCACCTGCTGTAACCGTC  
GATGTTGGTGCCAGGTGCCACCTGGGAAGGATGCTGTGCAGGGGGCTTGCCAACTTTGGTGGGTTTCAGAAGCCCCAG  
GCACTTGTGGCAGGCACAATTACAGCCCTCCCCAAAGATGCCACGTCCTTCTCCTGGAACCTGTGAATGTGTACCCG  
10 CAAGGCAGAGGCTGGTGAAGGCTGCAGGTGGAATCACGGCTGCCAGTCAGCCGATCTTAAGGTCATCCTGGATTATCTGG  
TGGGCCTGATATGGCCACAAGGGTCCCTAGAAGTGAGAGAGGGAGGCAGGGGAGAGTCAGAGAGGGGACGTGAGAAGGAC  
CACTGGCCACTGCTGGCTTTGAGATGGAGGAGGGGTCCCCAGCCAAGGAATGGGGGCAGCCGCTCCATGCTGGAAAAGC  
AAGCAATCCTCCCCGGTCTGAGGGCACACGGCCCTGCCACGCCTCGATTTAGGCCAGTGGGACCTGTTTCAGCTTTC  
CGGCCTCCAGAGCTGTAAGATGATGCGTTTGTTTCAGCCACTAAGCTGCAGTGATTGTCACAGCAGCAAAATGGAATAG  
15 CAGTACAGGGAATGAATACAGGGACAGTTCTCAGAGTGACTCTCAGCCCACCCCTGGG

Die Charakterisierung der Exons zeigte interessanterweise, daß die in unserer Patentanmeldung PCT/EP/98/03469 beschriebenen, funktionell wichtigen hTC-Protein-Domänen auf separaten Exons angeordnet sind. Das Telomerase-charakteristische T-Motiv befindet sich auf Exon 3. Die für die katalytische Funktion der Telomerase wichtigen RT (Reverse-Transkriptase)-Motive 1-7 liegen auf folgenden Exons: RT Motiv 1 und 2 auf Exon 4, RT Motiv 4 auf Exon 9, RT Motiv 5 auf Exon 10, RT Motiv 6 und 7 auf Exon 11. RT Motiv 3 liegt verteilt auf Exon 5 und 6 vor (s. Fig. 8).

10

Die Aufklärung der Exon-Intron-Struktur des hTC-Gens zeigt auch, daß die in unserer Patentanmeldung PCT/EP/98/03469 beschriebenen vier Deletions- bzw. Insertions-Varianten der hTC-cDNA ebenso wie drei weitere, in der Literatur (Kilian et al., 1997) beschriebene hTC-Insertions-Varianten höchstwahrscheinlich alternative Splice-Produkte darstellen. Wie in Fig. 8 gezeigt, lassen sich die Splice Varianten in zwei Gruppen einteilen: Deletionsvarianten und Insertionsvarianten.

15

Den hTC-Varianten der Deletionsgruppe fehlen spezifische Sequenzabschnitte. Die 36 bp in frame Deletion in Variante DEL1 resultiert höchstwahrscheinlich aus der Benutzung einer alternativen 3'-Splice Akzeptorsequenz in Exon 6, wodurch ein Teil des RT Motivs 3 verlorenght. In Variante DEL2 werden die normalen 5'-Splice Donor- und 3'-Splice-Akzeptor Sequenzen von Intron 6, 7 und 8 nicht benutzt. Stattdessen wird Exon 6 direkt an Exon 9 fusioniert, wodurch eine Verschiebung des offenen Leserahmens entsteht und in Exon 10 ein Stopcodon auftritt. Variante Del3 stellt eine Kombination aus Variante 1 und 2 dar.

20

25

Die Gruppe der Insertions-Varianten zeichnet sich durch die Insertion von Intronsequenzen aus, die zu vorzeitigen Translationsstop führen. Anstelle der normalerweise benutzten 5'-Splice Donorsequenz von Intron 5 wird eine alternative, 3'-lokalisierte Splicestelle in Variante INS1 benutzt, wodurch eine Insertion der ersten 38 bp aus Intron 4 zwischen Exon 4 und Exon 5 entsteht. Ebenso resultiert die

30

Insertion eines Intron 11-Sequenzbereichs in Variante INS2 aus der Benutzung einer alternativen 5'-Splice Donorsequenz in Intron 11. Da diese Variante in der Literatur (Kilian et al., 1997) nur unzureichend beschrieben wurde, läßt sich die genaue alternative 5'-Splice Donorsequenz dieser Variante nicht bestimmen. Die Insertion von Intron 14 Sequenzen zwischen Exon 14 und Exon 15 in Variante INS3 entsteht durch die Benutzung von einer alternativen 3'-Splice Akzeptorsequenz, wodurch der 3'-Teil von Intron 14 nicht gesplitt wird.

Die in unserer Patentanmeldung PCT/EP/98/03469 beschriebene hTC-Variante INS4 (Variante 4) zeichnet sich durch den Ersatz von Exon 15 und dem 5'-Teilbereich von Exon 16 durch die ersten 600 bp des Introns 14 aus. Diese Variante ist auf den Gebrauch einer alternativer internen 5'-Splice Donorsequenz in Intron 14 und einer alternativen 3'-Splice Akzeptorsequenz in Exon 16 zurückzuführen, woraus ein veränderter C-Terminus resultiert.

Die *in vivo*-Generation wahrscheinlich nicht-funktioneller hTC-Proteinvarianten, die mit der Funktion des vollständigen hTC-Proteins interferieren könnten, stellt zusätzlich zur Transkriptionsregulation einen möglichen Mechanismus dar, um die hTC-Proteinfunktion zu kontrollieren. Bis heute ist die Funktion der hTC-Splicevarianten nicht bekannt. Obwohl die meisten dieser Varianten vermutlich für Proteine ohne Reverse-Transkriptase-Aktivität kodieren, könnten sie dennoch eine entscheidende Rolle als transdominant-negative Telomerase-Regulatoren spielen, indem sie z.B. um die Interaktion mit wichtigen Bindungspartnern kompetieren.

Die Suche nach möglichen Transkriptionsfaktorbindungsstellen wurde mit dem „Find Pattern“-Algorithmus aus dem „GCG Sequenz Analysis“ Programmpaket der „Genetics Computer Group“ (Madison, USA) durchgeführt. Dadurch wurden verschiedene potentielle Bindungsstellen für Transkriptionsfaktoren in der Nukleotidsequenz von Intron 2 identifiziert, die in der Tab. 2 aufgelistet sind. Darüberhinaus wurde im Intron 1 eine Sp1-Bindungsstelle (Pos. 43) und im 5'-

untranslatiertem Bereich eine c-Myc-Bindungsstelle (cDNA-Position 29-34, vergl. Fig. 6) gefunden.

### Beispiel 6

5

Um den oder die Startpunkt(e) der hTC-Transkription in HL 60 Zellen zu ermitteln, wurde das 5'-Ende der hTC-mRNA durch Primer-Extension-Analyse bestimmt.

Es wurden 2 µg PolyA<sup>+</sup>-RNA aus HL-60-Zellen für 10 min bei 65°C denaturiert. Zur  
10 Primeranlagerung wurden 1 µl RNasin (30-40 U/ml) und 0,3-1 pmol radioaktiv  
markierter Primer (5'GTAAAGTTGTAGCTTACACTGGTTCTC 3'; 2,5-8x10<sup>5</sup>  
cpm) zugegeben und für 30 min bei 37°C in einem Gesamtvolumen von 20 µl  
inkubiert. Nach Zugabe von 10 µl 5xReverse Transkriptase-Puffer (Fa. Gibco-BRL),  
2 µl 10 mM dNTPs, 2 µl RNasin (s.o.), 5µl 0,1 M DTT (Fa. Gibco-BRL) 2 µl  
15 ThermoScript RT (15 U/µl; Fa. Gibco-BRL) und 9 µl DEPC-behandeltes Wasser  
erfolgte die Primer-Verlängerung in einem Gesamtvolumen für 1 h bei 58°C. Die  
Reaktion wurde durch 4 µl 0,5 M EDTA, pH 8,0, gestoppt und die RNA nach  
Zugabe von 1 µl RNaseA (10 mg/ml) für 30 min bei 37°C abgebaut. Hierauf wurden  
2,5 µg gescherte Kalbsthymus-DNA und 100 µl TE addiert und einmal mit 150µl  
20 Phenol/Chloroform (1:1) extrahiert. Die DNA wurde unter Zusatz von 15 µl 3 M Na-  
Acetat und 450 µl Ethanol für 45 min bei -70°C gefällt und anschließend für 15 min  
bei 14000 Upm abzentrifugiert. Das Präzipitat wurde einmal mit 70 %igem Ethanol  
gewaschen, luftgetrocknet und in 8 µl Sequenzierungs-Stopplösung gelöst. Nach 5  
min Denaturierung bei 80°C wurden die Proben auf ein 6 %iges Polyacrylamidgel  
25 aufgetragen und elektrophoretisch (Ausubel et al., 1987) aufgetrennt (Fig. 5).

Hierbei wurde eine Haupt-Transkriptionsstartstelle identifiziert, die 1767 bp 5' vom  
ATG-Startcodon der hTC-cDNA Sequenz lokalisiert ist (Nukleotidposition 3346 in  
Fig. 4). Die Nukleotidsequenz um diesen Haupttranskriptionsstart (TTA<sub>n</sub>TTGT)  
30 repräsentiert darüberhinaus ein Initiator-Element (Inr), das in 6 von 7 Nukleotiden

mit dem Konsensusmotiv (PyPyA<sub>n</sub>Na/tPyPy) (Smale, 1997) eines Initiator-Elementes übereinstimmt.

5 In unmittelbarer Nähe des experimentell identifizierten Haupt-Transkriptionsstartes konnte keine eindeutige TATA-Box identifiziert werden, so daß der hTC-Promoter wahrscheinlich in die Familie der TATA-losen Promotoren (Smale, 1997) einzuordnen ist. Allerdings wurde durch Bioinformatik Analyse eine potentielle TATA-Box von Nukleotidposition 1306 bis 1311 (Fig. 4) gefunden. Die zusätzlich um den Haupt-Transkriptionsstart beobachteten Neben-Transkriptionsstarts wurden auch bei  
10 anderen TATA-losen Promotoren beschrieben (Geng and Johnson, 1993), wie z.B. in den stark regulierten Promotoren einiger Zellzyklusgene (Wick *et al.*, 1995).

#### Beispiel 7

15 Zusätzlich zu dem in Beispiel 6 beschriebenen, in HL60 Zellen identifizierten Startpunkt des hTC Transkriptes, wurde ein weiterer Transkriptionsstartbereich in HL60 Zellen identifiziert. Anhand von RT-PCR-Analysen wurde die Region des Transkriptionsstarts des hTC-Gens in HL60 Zellen auf die bp -60 bis -105 eingegrenzt.

20 Unter Einsatz von 0,4 µg Poly A-RNA aus HL60 Zellen (Clontech) und dem genspezifischen Primer GSP13 (5'-CCTCCAAAGAGGTGGCTTCTTCGGC-3', cDNA-Position 920-897) wurde hierfür die cDNA mit Hilfe des „First Strand cDNA-Synthesis Kit“ (Clontech) nach Angaben der Hersteller synthetisiert. In einem  
25 Endvolumen von 50 µl wurden 1 µl cDNA mit 10 pmol dNTP-Mix versetzt und in 1xPCR-Reaktionspuffer F (PCR-Optimizer Kit der Fa. InVitrogen) und einem Unit Platinum-Taq-DNA Polymerase (Fa. Gibco/BRL) eine PCR-Reaktion durchgeführt. Als Primer wurden jeweils 10 pmol der nachfolgend definierten 5'- und 3'-Primer  
30 zugefügt. Die PCR wurde in 3 Schritten durchgeführt. An eine zweiminütige Denaturierung bei 94°C schlossen sich 36 PCR-Zyklen an, in denen die DNA zunächst für 45 sec bei 94°C denaturiert wurde und anschließend für 5 min bei 68°C die Primer

angelagert und die DNA-Kette verlängert wurde. Zum Abschluß folgte für 10 min eine Kettenverlängerung bei 68°C. Insgesamt wurden sechs verschiedene 5'-PCR Primer (Primer HTRT5B: 5'-CGCAGCCACTACCGCGAGGTGC-3', cDNA-Position 105 bis 126; Primer C5S: 5'-CTGCGTCCTGCTGCGCACGTGGGAAGC-3', 5'-flankierende Region -49 bis -23; Primer PRO-TEST1: 5'-CTCGCGGCGCGAGTTTCAGGCAG-3', 5'-flankierende Region -74 bis -52; Primer PRO-TEST2: 5'-CCAGCCCCTCCCCTTCCTTTCC-3', 5'-flankierende Region -112 bis -91; Primer PRO-TEST4: 5'-CCAGCTCCGCCTCCTCCGCGC-3', 5'-flankierende Region -191 - -171; Primer RP-3A: 5'-CTAGGCCGATTCGACCTCTCTCC-3', 5'-flankierende Region -427 bis -405) mit dem 3'-PCR Primer C5Rück (5'-GTCCCAGGGCACGCACACCAG-3', cDNA-Position 245 bis 225) kombiniert. Als Kontrolle wurde zusätzlich zu den Oligo-dT- und GSP13-geprimten cDNAs auch genomische DNA für die PCR eingesetzt. Wie in Fig. 9 gezeigt, wurde nur mit den Primerkombinationen HTRT5B-C5Rück, C5S-C5Rück und PRO-TEST1-C5Rück ein PCR-Produkt erhalten, was darauf hinweist, daß der Startpunkt der hTC-Transkription in der Region zwischen bp-60 und bp-105 liegt.

### Beispiel 8

In der ca. 11,2 kb isolierten 5'-flankierenden Region des hTC-Gens befinden sich mehrere extrem GC-reiche Bereiche, sog. CpG Islands. Ein CpG Islands mit einem GC-Gehalt von > 70 % reicht von bp - 1214 bis in Intron 2. Zwei weitere GC-reiche Bereiche mit einem GC-Gehalt von > 60 % reichen von bp -3872 bis bp -3113 bzw. bp -5363 bis bp -3941. Die Lage der CpG Islands ist in der Fig. 11 graphisch dargestellt.

Die Suche nach möglichen Transkriptionsfaktorbindungsstellen wurde mit dem „Find Pattern“-Algorithmus aus dem „GCG Sequenz Analysis“ Programmpaket der „Genetics Computer Group“ (Madison, USA) durchgeführt. Dadurch wurden verschiedene potentielle Bindungsstellen in der Region bis -900 bp upstream vom

Translations-Startcodon ATG indentifiziert: fünf Spl-Bindungsstellen, eine c-Myc-Bindungsstelle, eine CCAC-Box (Fig. 10). Zusätzlich wurden eine CCAAT-Box und eine zweite c-Myc-Bindungsstelle an den Positionen -1788 bzw. -3995 der 5'-flankierenden Region gefunden.

5

### Beispiel 9

Um die Aktivität des hTC-Promotors zu analysieren, wurden durch PCR-Amplifikation vier verschieden lange hTC-Promotorsequenzabschnitte generiert und 5' vor das Reportergen Luziferase in den Vektor pGL2 der Fa. Promega kloniert. Als DNA-Quelle für die PCR-Amplifikation wurde das aus dem Phagenklon P12 subklonierte, 8,5 kb große SacI-Fragment gewählt. In einem Endvolumen von 50 µl wurden 35 ng dieser DNA mit 10 pmol dNTP-Mix versetzt und in 1xPCR-Reaktionspuffer (PCR-Optimizer Kit der Fa. InVitrogen) und einem Unit Platinum-Taq-DNA Polymerase (Fa. Gibco/BRL) eine PCR-Reaktion durchgeführt. Als Primer wurden jeweils 20 pmol der nachfolgend definierten 5'- und 3'-Primer zugefügt. Die PCR wurde in 3 Schritten durchgeführt. An eine zweiminütige Denaturierung bei 94°C schlossen sich 30 PCR-Zyklen an, in denen die DNA zunächst für 45 sec bei 94°C denaturiert wurde und anschließend für 5 min bei 68°C die Primer angelagert und die DNA-Kette verlängert wurde. Zum Abschluß folgte für 10 min eine Kettenverlängerung bei 68°C. Als 3'-PCR-Primer wurde jeweils der Primer PK-3A (5'-GCAAGCTTGACGCAGCGCTGCCTGAAACTCG-3', Position -43 bis -65) gewählt, der einen Sequenzbereich 42 bp upstream vom START-Codon ATG erkennt. Durch Kombination des PK-3A-Primers mit dem 5'-PCR-Primer PK-5B (5'-CCAGATCTCTGGAACACAGAGTGGCAGTTTCC-3', Position -4093 bis -4070) wurde ein 4051 bp großes Promotor-Fragment amplifiziert (NPK8). Die Kombination des Primerpaares PK-3A und PK-5C (5'-CCAGATCTGCATGAAGTGTGTGGGATTTCAG-3', Position -3120 bis -3096) führte zur Amplifikation eines 3078 bp großen Promotorfragmentes (NPK15). Ein 2068 bp großes Promotorfragment wurde durch die Verwendung der Primerkombination PK-3A und PK-5D (5'-

GGAGATCTGATCTTGGCTTACTGCAGCCTCTG-3', Position -2110 bis -2087) amplifiziert (NPK22). Der Einsatz der Primerkombination PK-3A und PK-5E (5'-GGAGATCTGTCTGGATTCCTGGGAAGTCCTCA-3', Position -1125 bis -1102) führte schließlich zur Amplifikation eines 1083 bp großen Promotorfragmentes (NPK27). Der PK-3A Primer enthält eine HindIII Erkennungssequenz. Die verschiedenen 5'-Primer enthalten eine BglII-Erkennungssequenz.

Die entstandenen PCR-Produkte wurden mit Hilfe des QIA quick spin PCR Purification Kits der Fa. Qiagen nach Angaben der Hersteller aufgereinigt und anschließend mit den Restriktionsenzymen BglII und HindIII verdaut. Mit den gleichen Restriktionsenzymen wurde der pGL2-Promotor-Vektor verdaut und der in diesem Vektor enthaltene SV40-Promotor freigesetzt und abgetrennt. Die PCR-Promotorfragmente wurden in den Vektor ligiert, in kompetente DH5 $\alpha$ -Bakterien der Fa. Gibco/BRL transformiert. Aus transformierten Bakterienklonen wurde DNA für die nachfolgend beschriebenen Promotor-Aktivitäts-Analysen mit Hilfe des Qiagen Plasmid-Kits der Fa. Qiagen isoliert.

#### Beispiel 10

Die Aktivität des hTC-Promotors wurde in transienten Transfektionen in eukaryotischen Zellen analysiert.

Alle Arbeiten mit eukaryotischen Zellen erfolgten an einem sterilen Arbeitsplatz. CHO-K1 und HEK 293 Zellen wurden von der American Type Culture collection bezogen.

CHO-K1 Zellen wurden in DMEM Nut Mix F-12 Zellkulturmedium (Fa. Gibco-BRL, Bestellnummer: 21331-020) mit 0,15 % Streptomycin/Penezillin, 2 mM Glutamin und 10 % FCS (Fa. Gibco-BRL) gehalten.

HEK 293 Zellen wurden in DMOD Zellkulturmedium (Fa. Gibco-BRL, Bestellnummer: 41965-039) mit 0,15 % Streptomycin/Penicillin, 2 mM Glutamin und 10 % FCS (Fa. Gibco-BRL) kultiviert.

- 5 CHO-K1 und HEK 293 Zellen wurden in wasssergesättigter Atmosphäre bei 37°C unter Begasung mit 5 % CO<sub>2</sub> kultiviert. Bei konfluentem Zellrasen wurde das Medium abgesaugt, die Zellen mit PBS (100 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7,2; 150 mM NaCl) gewaschen und durch Zugabe einer Trypsin-EDTA Lösung (Fa. Gibco-BRL) abgelöst. Das Trypsin wurde durch Mediumzugabe inaktiviert und die Zellzahl mit  
10 einer Neubauer-Zählkammer ermittelt, um die Zellen in gewünschter Dichte auszu-plattieren.

- Für die Transfektion wurden pro Well jeweils 2x 10<sup>5</sup> –HEK 293 Zellen in einer 24-well Zellkulturplatte ausplattiert. Nach 3 Stunden wurde das HEK 293 Medium  
15 entfernt. Für die Transfektion wurden bis zu 2,5 µg Plasmid-DNA, 1 µg eines CMV β-Gal Plasmidkonstruktes (Fa. Stratagene, Bestellnummer: 200388), 200 µl serum-freies Medium und 10 µl Transfektionsreagenz (DOTAP der Fa. Boehringer Mannheim) für 15 Minuten bei Raumtemperatur inkubiert und anschließend auf die HEK 293 Zellen gleichmäßig aufgetropft. Nach 3 Stunden wurden 1,5 ml Medium  
20 hinzugegeben. Nach 20 Stunden wurde das Medium gewechselt. Nach weiteren 24 Stunden wurden die Zellen zur Bestimmung der Luziferase- und der β-Gal-Aktivität geerntet. Dazu wurden die Zellen im Zellkultur-Lysisreagenz (25 mM Tris [pH 7,8] mit H<sub>3</sub>PO<sub>4</sub>; 2 mM CDTA; 2 mM DTT; 10% Glycerol; 1% Triton X-100) für 15 Minuten bei Raumtemperatur lysiert. Zwanzig µl dieses Zellysats wurden mit 100 µl  
25 Luziferase-Assaypuffer (20 mM Tricin; 1,07 mM (MgCO<sub>3</sub>)<sub>4</sub> Mg(OH)<sub>2</sub>·5H<sub>2</sub>O; 2,67 mM MgSO<sub>4</sub>; 0,1 mM EDTA; 33,3 mM DTT; 270 µM Coenzym A; 470 µM Luciferin, 530 µM ATP) gemischt und das durch die Luziferase generierte Licht gemessen.

- 30 Zur Messung der β-Galaktosidaseaktivität wurden gleiche Mengen Zellysat und β-Galaktosidase-Assaypuffer (100 mM Natriumphosphatpuffer pH 7,3; 1 mM MgCl<sub>2</sub>;

50 mM  $\beta$ -Merkaptoethanol; 0,665 mg/ml ONPG) für mindestens 30 Minuten bei 37°C oder bis eine leichte Gelbfärbung auftrat, inkubiert. Die Reaktion wurde durch Zugabe von 100  $\mu$ l 1 M  $\text{Na}_2\text{CO}_3$  gestoppt und die Absorption bei 420 nm bestimmt.

- 5 Für die Analyse des hTC-Promotors wurden vier verschieden lange hTC-Promotorsequenzabschnitte 5' vor das Reportergen Luziferase kloniert (vergl. Beispiel 9).

In der Fig. 11 sind die relativen Luziferase Aktivitäten zweier unabhängiger Transfektionen mit den Konstrukten NPK8, NPK15, NPK22 und NPK27 in HEK  
10 293 Zellen aufgetragen. Jedes Experiment wurde in Duplikaten durchgeführt. Darüberhinaus wurde die Standardabweichung angegeben. Das Konstrukt NPK 27 zeigt eine 40fach höhere Luziferaseaktivität als die Basalaktivität des promotorlosen Luziferase-Kontrollkonstrutes (pGL2-basic) und eine 2 bis 3fach höhere Aktivität als das SV40 Promotorkontroll-Konstrukt (pGL2PRO). Interessanterweise wurde im  
15 Vergleich zu dem Konstrukt NPK27 eine 2 bis 3fach geringere Luziferaseaktivität in mit längeren hTC Promotorkonstrukten (NPK8, NPK15, NPK22) transfizierten Zellen beobachtet. Ähnliche Ergebnisse wurden auch in CHO Zellen beobachtet (Daten nicht gezeigt).

## Literaturverzeichnis

- 5 Allsopp, R. C., Vazire, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, A.B., Greider, C.W. und Harley, C.B. (1992). Telomere length predicts replicative capacity of human fibroblasts. Proc. Natl. Acad. Sci. 89, 10114-10118.
- 10 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987). Current protocols in molecular biology. Greene Publishing Associates and Wiley-Intersciences, New York.
- Blasco, M. A., Rizen, M., Greider, C. W. und Hanahan, D. (1996). Differential regulation of telomerase activity and telomerase RNA during multistage tumorigenesis. Nature Genetics 12, 200-204.
- 15 Broccoli, D., Young, J. W. und deLange, T. (1995). Telomerase activity in normal and malignant hematopoietic cells. Proc. Natl. Acad. Sci. 92, 9082-9086.
- Counter, C. M., Avilion, A. A., LeFeuvre, C. E., Stewart, N. G. Greider, C.W. Harley, C. B. und Bacchetti S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J. 11, 1921-1929.
- 20 Feng, J., Funk, W. D., Wang, S.-S., Weinrich, S. L., Avilion, A.A., Chiu, C.-P., Adams, R.R., Chang, E., Allsopp, R.C., Yu, J., Le, S., West, M.D., Harley, C.B., Andrews, W.H., Greider, C.W. und Villeponteau, B. (1995). The RNA component of human telomerase. Science 269, 1236-1241.
- 25 Geng, Y., and Johnson, L.F. (1993). Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA less mouse thymidine synthase promoter. Mol. Cell. Biol 14:4894.
- 30 Goldstein, S. (1990). Replicative senescence: The human fibroblast comes of age. Science 249, 1129-1133.
- 35 Harley, C.B., Futcher, A.B., Greider, C.W., 1990. Telomeres shorten during ageing of human fibroblasts. Nature 345, 458-460.

- Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D.K. und Allshire, R.C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. *Nature* 346, 866-868.
- 5     Hiyaama, K., Hirai, Y., Kyoizumi, S., Akiyama, M., Hiyaama, E., Piatyszek, M.A., Shay, J.W., Ishioka, S. und Yamakido, M. (1995). Activation of telomerase in human lymphocytes and hematopoietic progenitor cells. *J. Immunol.* 155, 3711-3715.
- 10     Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C. B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L. und Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011-2015.
- Latchman, D.S. (1991). Eukaryotic transcription factors. Academic Press Limited, London.
- 15     Lingner, J., Hughes, T.R., Shevchenko, A., Mann, M., Lundblad, V. und Cech T.R. (1997). Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276: 561-567.
- Lundblad, V. und Szostak, J. W. (1989). A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57, 633-643.
- 20     McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics* 26, 234-282.
- 25     Meyne, J., Ratliff, R. L. und Moyzis, R. K. (1989). Conservation of the human telomere sequence (TTAGGG)<sub>n</sub> among vertebrates. *Proc. Natl. Acad. Sci.* 86, 7049-7053.
- Olovnikov, A. M. (1973). A theory of marginotomy. *J. Theor. Biol.* 41, 181-190.
- 30     Sandell, L. L. und Zakian, V. A. (1993). Loss of a yeast telomere: Arrest, recovery and chromosome loss. *Cell* 75, 729-739.
- Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* 15, 7155-7174.
- 35     Smale, S.T. and Baltimore, D. (1989). The „initiator“ as a transcription control element. *Cell* 57:103-113.

Smale, S.T. (1997). Transcription initiation from TATA-less promoters within eukaryotic protein-coding genes. *Biochimica et Biophysica Acta* 1351, 73-88.

5 Shay, J. W. (1997). Telomerae and Cancer. Ciba Foundation Meeting: Telomeres and Telomerase. London.

Vaziri, H., Dragowska, W., Allsopp, R. C., Thomas, T. E., Harley, C.B. und Landsdorp, P.M. (1994). Evidence for a mitotic clock in human hematopoietic stem cells: Loss of telomeric DNA with  
10 age. *Proc. Natl. Acad. Sci.* 91, 9857-9860.

Wick, M., Härönen, R., Mumberg, D., Bürger, C., Olsen, B.R., Budarf, M.L., Apte, S. S. and Müller, R. (1995). Structure of the human TIMP-3 gene and its cell-cycle-regulated promoter. *Biochemical Journal* 311, 549-554.  
15

Zakian, V. A. (1995). Telomeres: Beginning to understand the end. *Science* 270, 1601-1607.

**Patentansprüche**

1. Regulatorische DNA-Sequenzen für das Gen der humanen katalytischen  
Telomerase-Untereinheit.  
5
2. DNA-Sequenzen gemäß Anspruch 1, dadurch gekennzeichnet, daß es sich um  
Intronsequenzen gemäß SEQ ID NO 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16,  
17, 18, 19 und/oder 20 oder um regulatorisch wirksame Fragmente dieser  
Sequenzen handelt.  
10
3. DNA-Sequenzen gemäß Anspruch 1, dadurch gekennzeichnet, daß es sich um  
die 5'-flankierende regulatorische DNA-Sequenz für das Gen der humanen  
katalytischen Telomerase-Untereinheit gemäß Fig. 10 (SEQ ID NO 3) oder  
um regulatorisch wirksame Fragmente dieser DNA-Sequenz handelt.  
15
4. Rekombinantes Konstrukt, enthaltend eine DNA-Sequenz gemäß einem der  
Ansprüche 1 bis 3.
5. Rekombinantes Konstrukt gemäß Anspruch 4, dadurch gekennzeichnet, daß  
es weiterhin eine oder mehrere DNA-Sequenzen enthält, die für Polypeptide  
oder Proteine kodieren.  
20
6. Vektor, enthaltend ein rekombinantes Konstrukt gemäß Anspruch 4 oder 5.
- 25 7. Verwendung von rekombinanten Konstrukten bzw. Vektoren gemäß einem  
der Ansprüche 4 bis 6 zur Herstellung von Arzneimitteln.
8. Rekombinante Wirtszellen, enthaltend rekombinante Konstrukte bzw. Vek-  
toren gemäß einem der Ansprüche 4 bis 6.

9. Verfahren zur Identifizierung von Substanzen, die die Promotor-, Silencer- oder Enhanceraktivität der humanen katalytischen Telomerase-Untereinheit beeinflussen, das folgende Schritte umfaßt:
- 5           A. Zugabe einer Kandidatensubstanz zu einer Wirtszelle, enthaltend DNA-Sequenzen gemäß einem der Ansprüche 1 bis 3, funktionell verknüpft mit einem Reporter-gen,
- B. Messung des Substanzeffektes auf die Reporter-genexpression.
- 10
10. Verfahren zur Identifizierung von Faktoren, die spezifisch an die DNA gemäß einem der Ansprüche 1 bis 3 oder an Fragmente davon binden, dadurch gekennzeichnet, daß man eine Expressions-cDNA-Bibliothek mit einer DNA-Sequenz gemäß einem der Ansprüche 1 bis 3 oder Teilfragmenten unterschiedlichster Länge als Sonde screent.
- 15
11. Transgene Tiere, enthaltend rekombinante Konstrukte bzw. Vektoren gemäß Ansprüchen 4 bis 6.
- 20
12. Verfahren zur Detektion Telomerase-assoziiierter Zustände bei einem Patienten, das folgende Schritte umfaßt:
- A. Inkubation eines rekombinanten Konstruktes bzw. Vektors gemäß Ansprüchen 4 bis 6 das bzw. der zusätzlich ein Reporter-gen enthält
- 25           mit Körperflüssigkeiten oder zellulären Proben,
- B. Detektion der Reporter-genaktivität, um einen diagnostischen Wert zu erhalten,

- C. Vergleich des diagnostischen Wertes mit Standardwerten für das Reportergenkonstrukt in standardisierten normalen Zellen oder Körperflüssigkeiten des gleichen Typs wie die Testprobe.

Fig. 1

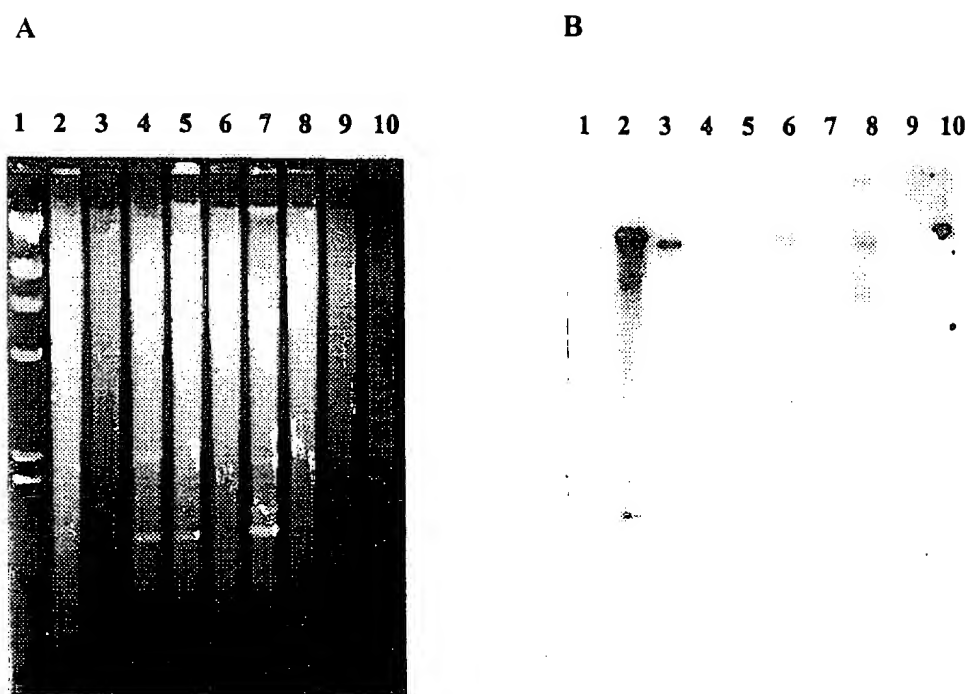


Fig. 2

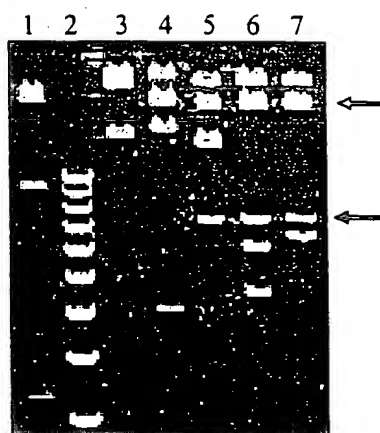


Fig. 3

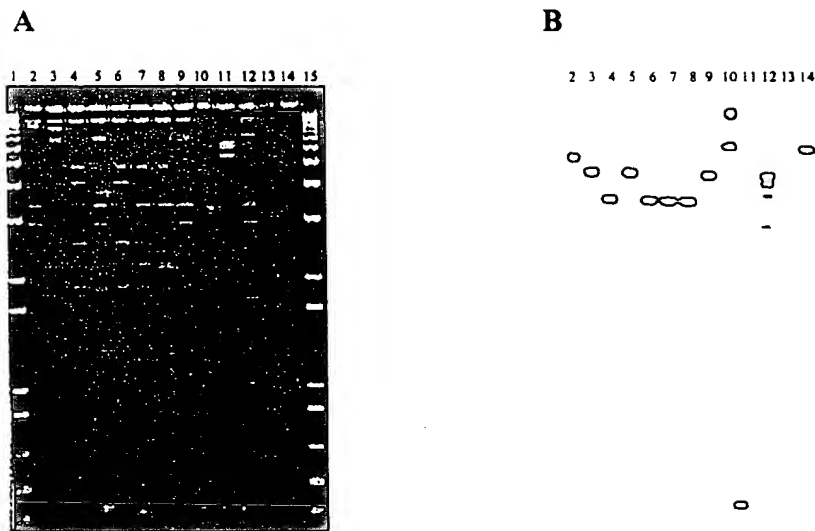


Fig. 4

GAGCTCTGAA	CCGTGGAAAC	GAACATGACC	CTTGCCCTGCC	TGCTTCCCTG	GGTGGGTCAA	GGTAATGAA	70
GTGGTGTGCA	GGAAATGGCC	ATGTAAATTA	CACGACTCTG	CTGATGGGGA	CCGTTCCCTC	CATCATATT	140
CATCTTCACC	CCCAAGGACT	GAATGATTCC	AGCAACTTCT	TCGGGTGTGA	CAAGCCATGA	CAAACTCAG	210
TACAAACACC	ACTCTTTTAC	TAGGCCACCA	GAGCACGGGC	CACACCCCTG	ATATATTAA	AGTCCAGGAG	280
AGATGAGGCT	GCTTTCAGCC	ACCAGGCTGG	GGTGACAACA	GCGGCTGAAC	AGTCTGTTCC	TCTAGACTAG	350
TAGACCCCTGG	CAGGCACTCC	CCCAAATTCT	AGGGCCTGGT	TGCTGCTTCC	CGAGGGCGCC	ATCTGCCCTG	420
GAGACTCAGC	CTGGGGTGCC	ACACTGAGGC	CAGCCCTGTG	TCCACACCCT	CCGCCTCCAG	GCCTCAGCTT	490
CTCCAGCAGC	TTCTAAACC	CTGGGTGGGC	CGTGTTCAG	CGTACTGTG	TCACCTGTCC	CACTGTGTCT	560
TGCTCAGCG	ACGTAGCTCG	CACGGTTTCT	CCTCACATGG	GGTGTCTGTC	TCCTTCCCCA	ACACTCACAT	630
GCGTTGAAGG	GAGGAGATTC	TGCGCCTCCC	AGACTGGCTC	CTCTGAGCCT	GAACCTGGCT	CGTGGCCCCC	700
GATGCAGGTT	CCTGGCGTCC	GGCTGCACGC	TGACCTCCAT	TTCCAGGCGC	TCCCCGTCTC	CTGTCTATCTG	770
CCGGGGCCTG	CCGGTGTGTT	CTTCTGTTTC	TGTGCTCCTT	TCCACGTCCA	GCTGCGGTG	TCTCTGCCCC	840
CTAGGGTCTC	GGGGTTTTTA	TAGGCATAGG	ACGGGGGCGT	GGTGGGCCAG	GGCGCTCTTG	GGAAATGCAA	910
CATTGCGGTG	TGAAAGTAGG	AGTGCTGTG	CTCACCTAGG	TCCACGGGCA	CAGGCTCTGG	GATGGAGCCC	980
CCGCCAGGGA	CCCGCCCTTC	TCTGCCCAGC	ACTTCTCTGC	CCCCCTCCCT	CTGGAACACA	GAGTGGCAGT	1050
TTCCACAAGC	ACTAAGCATC	CTCTTCCCAA	AAGACCAGC	ATTGGCACCC	CTGGACATTT	GGCCCCAGC	1120
CTGGGGTAAT	CACGTGACTA	CGCACATCAT	GTACACACTC	CCGTCCACGA	CCGACCCCTG	CTGTTTATTT	1190
TTAATAGCTA	CAAAGCAGGG	AAATCCCTGC	TAAATGTCC	TTTAACAAAC	TGGTTAAACA	AACGGGTCCA	1260
TCCGCACGGT	GGACAGTTCC	TCACAGTGAA	GAGGAACATG	CCGTTTATAA	AGCCTGCAGG	CATCTCAAGG	1330
GAATTACGCT	GAGTCAAAAC	TGCCACCTCC	ATGGGATACG	TACGCAACAT	GCTCAAAAAG	AAAGAATTC	1400
ACCCCATGGC	AGGGGAGTGG	TTAGGGGGGT	TAAGGACGGT	GGGGGCGGCA	GCTGGGGGCT	ACTGCAAGCA	1470
CCTTTTACTA	AAGCCAGTTT	CCTGGTCTG	ATGGTATTGG	CTCAGTTATG	GGAGACTAAC	CATAGGGGAG	1540
TGGGGATGGG	GGAAACCCGA	GGCTGTGCCA	TCTTTGCCAT	GCCCGAGTGT	CCTGGGCAGG	ATAATGCTCT	1610
AGAGATGCCC	ACGTCTCTGAT	TCCCCCAAAC	CTGTGGACAG	AACCCGCCCG	GCCCCAGGCT	CTTTGCAGGT	1680
GTGATCTCCG	TGAGGACCCT	GAGGTCTGGG	ATCCTTCCGG	ACTACCTGCA	GGCCCCGAAA	GTAATCCAGG	1750
GGTTCTGGGA	AGAGCGGGCC	AGGAGGGTCA	GAGGGGGGCA	GCCTCAGGAC	GATGGAGGCA	GTCACTCTGA	1820
GCGTGAAAAG	GGAGGGAGGG	CCTCGAGCCC	AGGCCTGCAA	GCGCCTCCAG	AAGCTGGAAA	AAGCGGGGAA	1890
GGGACCCTCC	ACGGAGCCTG	CAGCAGGAAG	GCACGGCTGG	CCCTTAGCCC	ACCAGGGCCC	ATCGTGGACC	1960
TCCGGCCTCC	GTGCCATAGG	ACGGCACTCG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	ATTGTCAGAA	2030
GCAACAGGAA	ACCCATGCAC	TGTGAATCTA	GGATTATTTT	AAAACAAAGG	TTTACAGAAA	CATCCAAAGG	2100
CAGGGCTGAA	GTGCTCCCG	GCAAGGGCAG	GGCAGGCACG	AGTGATTTTA	TTAGCTATT	TTATTTTATT	2170
TACTTACTTT	CTGAGACAGA	GTTATGCTCT	TGTTGCCCAG	GCTGGAGTGC	AGCGGCATGA	TCTTGGCTCA	2240
GTGCAACCTC	CGTCTCCTGG	GTTCAAGCAA	TTCTCGTGCC	TCAGCCTCCC	AAGTAGCTGG	GATTTCAAGC	2310
CTGCCACACC	ACACCCGGCT	AATTTTGTAT	TTTAGTAGA	GATGGGCTTT	CACCATGTTG	GTCAAGCTGA	2380
TCTCAAAATC	CTGACCTCAG	GTGATCCGCC	CACCTCAGCC	TCCCAAAGTG	CTGGGATTAC	AGGCATGAGC	2450
CACTGCACCT	GGCCTATTTA	ACCATTTTAA	AACTTCCCTG	GGCTCAAGTC	ACACCCACTG	GTAGGAGGTT	2520
CATGGAGTTC	AATTTCCCTT	TACTCAGGA	GTTACCCTCC	TTTGATATTT	TCTGTAAATC	TTCCGTAGACT	2590
GGGGATACAC	CGTCTCTTGA	CATATTCACA	GTTTCTGTGA	CCACCTGTTA	TCCCATGGGA	TCCCATGCAG	2660
GGGCAGCTGG	GAGGCTGCAG	GCTTCAGGTC	CCAGTGGGGT	TGCCATCTGC	CAGTAGAAAC	CTGATGTAGA	2730
ATCAGGGCGC	AAGTGTGGAC	ACTGTCTCTG	ATCTCAATGT	CTCAGTGTGT	GCTGAACAT	GTAGAAATTA	2800
AAGTCCATCC	CTCCTACTCT	ACTGGGATTG	AGCCCTTCCC	CTATCCCCCC	CCAGGGGGCAG	AGGAGTTCTT	2870
CTCACTCCTG	TGGAGGAAGG	AATGATACTT	TGTTATTTTT	CACCTGCTGGT	ACTGAATCCA	CTGTTTCATT	2940
TGTTGGTTTG	TTTGTTTTGT	TTTGAGAGGC	GGTTTCACTC	TTGTGTCTCA	GGCTGGAGGG	AGTGCAATGG	3010
CGCGATCTTG	GCTTACTGCA	GCCTCTGCC	CCCAGGTTCA	AGTGATTCTC	CTGCTTCCGC	CTCCCATTTG	3080
GCTGGGATTA	CAGGCACCCG	CCACCATGCC	CAGCTAATTT	TTTGTATTTT	TAGTAGAGAC	GGGGGTGGGT	3150
GGGGTTCAAC	ATGTTGGCCA	GGCTGGTCTC	GAACCTCTGA	CCTCAGATGA	TCCACCTGCC	TCTGCCTCCT	3220
AAAGTGCTGG	GATTACAGGT	GTGAGCCACC	ATGCCAGCT	CAGAAATTTAC	TCTGTTTAGA	AACATCTGGG	3290
TCTGAGGTAG	GAAGCTCACC	CCACTCAAGT	GTTGTGGTGT	TTTAAGCCAA	TGATAGAATT	TTTTTATTGT	3360
TGTTAGAACCA	CTCTTGATGT	TTTACACTGT	GATGACTAAG	ACATCATCAG	CTTTTCAAAG	ACACACTAAC	3430
TGCACCCATA	ATACTGGGGT	GTCTTCTGGG	TATCAGCAAT	CTTCATTGAA	TGCCGGGAGG	CGTTTCTCTG	3500
CCATGCACAT	GGTGTTAATT	ACTCCAGCAT	AATCTTCTGC	TTCCATTTCT	TCTCTTCCCT	CTTTTAAAT	3570
TGTGTTTTCT	ATGTTGGCTT	CTCTGCAGAG	AACCAGTGTA	AGCTACAAC	TAACTTTTGT	TGGACAAAT	3640
TTTCCAAACC	GCCCCCTTGC	CCTAGTGGCA	GAGACAATTC	ACAAACACAG	CCCTTTAAAA	AGGCTTAGGG	3710
ATCACTAAGG	GGATTTCCTAG	AAGAGCCGAC	TGTAATCCTA	AGTATTTACA	AGACGAGGCT	AACCTCCAGC	3780
GAGCGTGACA	GCCCAGGGAG	GGTGCGAGGC	CTGTTCAAAT	GCTAGCTCCA	TAAATAAAGC	AAATTCCTCC	3850
GGCAGTTTCT	GAAAGTAGGA	AAGGTTACAT	TTAAGGTTGC	GTTTGTTAGC	ATTTCAAGTG	TTGCCGACCT	3920
CAGCTACAGC	ATCCCTGCAA	GGCCTCGGGA	GACCCAGAG	TTTCTCGCCC	CCTTAGATCC	AAACTTGAGC	3990
AACCCGGAGT	CTGGATTCTT	GGGAAGTCTT	CAGCTGTCTT	GCGGTTGTGC	CGGGGCCCCA	GGTCTGGAGG	4060
GGACCAGTGG	CCGTGTGGCT	TCTACTGCTG	GGCTGGAAGT	CGGGCCTCCT	AGCTCTGCAG	TCCGAGGCTT	4130
GGAGCCAGGT	GCCTGGACCC	CGAGGCTGCC	CTCCACCTCG	TGCGGGCGGG	ATGTGACCCG	ATGTTGGGCT	4200
CATCTGCCAG	ACAGAGTGCC	GGGGCCACAG	GTCAGGGCCG	TGTTGGCTGG	TGTGAGGCGC	CCGGTCCGCG	4270
GCCAGCAGGA	GCGCCTGGCT	CCATTTCCCA	CCCTTTCTCG	ACGGGACCCG	CCCGGTGGGT	GATTAACAGA	4340
TTTGGGGTGG	TTTGCTCATG	GTGGGGACCC	CTCGCCGCTT	GAGAACCTGC	AAAGAGAAAT	GACGGCCCTG	4410
TGTCAGGAG	CCCAAGTCCG	GGGGAAGTGT	TGCAGGGAGG	CACTCCGGGA	GGTCCCGCGT	GCCCGTCCAG	4480
GGAGCAATGC	GTCCCTCGGT	TGTTCCCGAG	CGCGCTCTAC	GCGCCTCCGT	CCTCCCTCTC	ACGTCGCGCA	4550
TTGCTGGTGC	CCGGAGCCCG	ACGCCCCCGG	TCCGGACCTG	GAGGCAGCCC	TGGGTCTCCG	GATCAGGCCA	4620
GCGGCCAAAG	GGTCCGCGCA	CGCACCTGTT	CCCAGGGCCT	CCACATCATG	GCCCCCTCCT	CGGGTTACCC	4690

Fig. 4 (Fortsetzung)

```

CACAGCCTAG GCCGATTCGA CCTCTCTCCG CTGGGGCCCT CGCTGGCGTC CCTGCACCCT GGGAGCGCGA 4760
GCGGCGCGCG GGGGGGAAG GCGGGCCAG ACCCCCGGT CCGCCCGGAG CAGCTGCGCT GTCGGGGCCA 4830
GGCCGGGCTC CAGTGGATT GCGGGGACA GACGCCAGG ACCGCGCTCC CCACGTGGCG GAGGGACTGG 4900
GGACCCGGGC ACCCGTCTG CCCCTTCACC TTCCAGCTCC GCCTCCTCCG GCGGACCCC GCCCGTCCC 4970
GACCCCTCCC GGGTCCCCGG CCCAGCCCC TCCGGGCCCT CCCAGCCCCT CCCCTTCCTT TCCGCGGCCC 5040
CGCCCTCTCC TCGCGGCGG AGTTTCAGG AGCGCTGCGT CCGCTGCGC ACCTGGGAAG CCCTGGCCCC 5110
GGCCACCCCC GCGATG 5126

```

Fig. 5

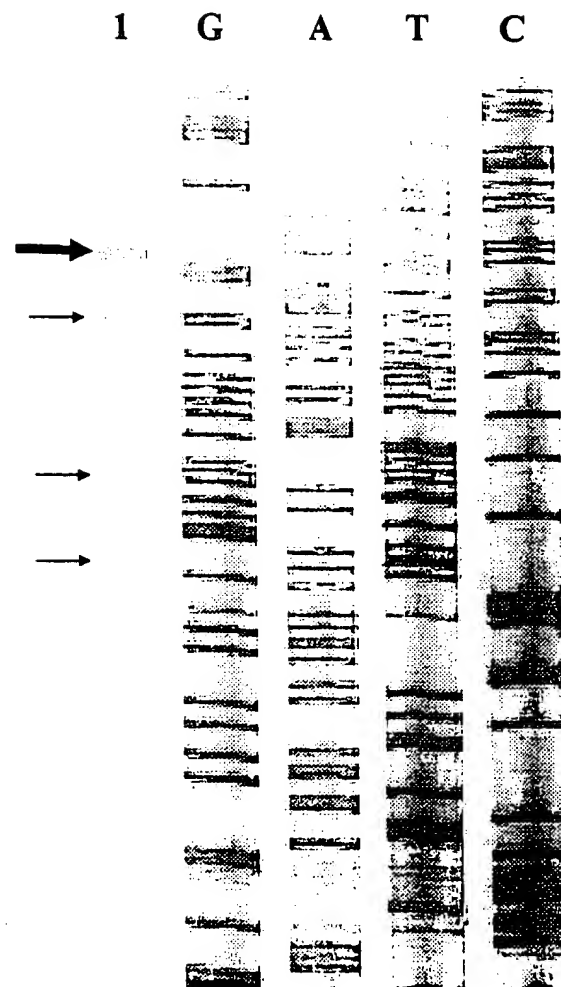


Fig. 6

GTTTCAGGCA	GCGCTGCGTC	CTGCTGCGCA	CGTGGGAAGC	CCTGGCCCCG	GCCACCCCCG	CGATGCCGCG	70
CGCTCCCCGC	TGCCGAGCCG	TGCGCTCCCT	GCTGCGCAGC	CACTACCCGC	AGGTGCTGCC	GCTGGCCACG	140
TTCGTGCGGC	GCTTGGGGCC	CCAGGGCTGG	CGGCTGGTGC	AGCGCGGGGA	CCCCGGGGCT	TTCCGCGCGC	210
TGGTGGCCCA	GTGCTTGGTG	TGCGTGCCCT	GGGACGCACG	GCCGCCCCCC	GCCGCCCCCT	CCTTCCGCCA	280
GGTGTCCCTGC	CTGAAGGAGC	TGGTGGCCCG	AGTGTGTGAG	AGGCTGTGCG	AGCGCGGGCG	GAAGAAGCTG	350
CTGGCCCTTCG	GCTTCCGCGT	GCTGGACGGG	GCCCGCGGGG	GCCCCCCCCA	GGCCTTCACC	ACCAGCGTGC	420
GCAGCTACCT	GCCCAACACG	GTGACCGACG	CACTGCGGGG	GAGCGGGGCG	TGGGGGCTGC	TGCTGCGCCG	490
CGTGGGCGAC	GACGTGCTGG	TTACCTGTCT	GGCAGCGTGC	GCGCTCTTTG	TGCTGGTGGC	TCCAGGTGCG	560
GCCTACACGG	TGTGCGGGCC	GCCGCTGTAC	CAGCTCGGCG	CTGCCACTCA	GGCCCCGGCC	CCGCCACACG	630
CTAGTGGACC	CCGAAGGCGT	CTGGGATGCG	AACGGGCTTG	GAACCATAGC	GTCAGGGAGG	CCGGGGTCCC	700
CCTGGGCTTG	CCAGCCCCGG	GTGCGAGGAG	GCGCGGGGGC	AGTGCCAGCC	GAAGTCTGCC	GTTGCCCAAG	770
AGGCCCAAGC	GTGGCGCTGC	CCCTGAGCCG	GAGCGGACGC	CCGTGGGGCA	GGGGTCTCTG	GCCCACCCGG	840
GCAGGACGCG	TGGACCGAGT	GACCGTGGTT	TCTGTGTGGT	GTCACCTGCC	AGACCCGCGG	AAGAAGCCAC	910
CTCTTTGGAG	GGTGCCTCT	CTGGCAGCGC	CCACTCCAC	CCATCCGTGG	GCCGCCACGA	CCAGCGGGGC	980
CCCCCATCCA	CATCGCGGCC	ACCACGTCCC	TGGGACACGC	CTTGTCCTCC	GGTGTACGCC	GAGACCAAGC	1050
ACTTCTCTTA	CTCCTCAGCG	GACAAGGAGC	AGCTGCGGCC	CTCCTTCTTA	CTCAGCTCTC	TGAGGCCCCAG	1120
CCTGACTGGC	GCTCGGAGGC	TGCTGGAGAC	CATCTTTCTG	GGTTCAGGCG	CCTGGATGCC	AGGGAATCCC	1190
CGCAGGTTGC	CCCGCTGTGC	CCAGCGCTAC	TGGCAAAATG	GGCCCTGTGT	TCTGGAGCTG	CTTGGGAACC	1260
ACGCGCAGTG	CCCTACGGCG	GTGCTCTCTA	AGACGCACTG	CCCGCTGCGA	GCTGCGGTCA	CCCCAGCAGC	1330
CGGTGTCTGT	GCCCCGGAGA	AGCCCCAGGG	CTCTGTGCGG	GCCCCCGAGG	AGGAGGACAC	AGACCCCCGT	1400
CGCCTGGTGC	AGCTGTCTCG	CCAGCACAGC	AGCCCTTGGC	AGGTGTACGG	CTTCGTGCGG	GCCTGCCTGC	1470
GCCGCGCTGT	GCCCCCAGGC	CTCTGGGGCT	CCAGGCACAA	CGAACGCCGC	TTCTCTAGGA	ACACCAAGAA	1540
GTTCTATCTC	CTGGGGGAAGC	ATGCCAAGCT	CTCGCTGCAG	GAGCTGACGT	GGAAGATGAG	CGTCCGGGAC	1610
TGCGGTTGGC	TGCGCAGGAG	CCCAGGGGTT	GGCTGTGTTC	CGGCCGCAGA	GCACCGTCTG	CGTGAGGAGA	1680
TCTTGGCCAA	GTTCTGTGAC	TGGCTGATGA	GTGTGTACGT	CGTGCAGCTG	CTCAGGTCTT	TCTTTTATGT	1750
CACGGAGACC	ACGTTTCAAA	AGAACAGGCT	CTTTTCTTAC	CGGAAGAGTG	TCTGGAGCAA	GTTCCAAAGC	1820
ATTGGAATCA	GACAGCACTT	GAAGAGGGTG	CAGCTGCGGG	AGCTGTGCGA	AGCAGAGGTC	AGGCAGCATC	1890
GGGAAGCCAG	GCCCCGCCCTG	CTGACGTCCA	GACTCCGCTT	CATCCCCAAG	CCTGACGGGC	TGCGGCCGAT	1960
TGTGAACATG	GACTACGTGC	TGGGAGCCAG	AACGTTCCGC	AGAGAAAAGA	GGGCCGAGCG	TCTCACCTCG	2030
AGGGTGAAGG	CACCTGTTAC	CGTGTCTCAAC	TACGAGCGGG	CGCGCGCGCC	CGGCTCTCTG	GGCGCCTCTG	2100
TGCTGGGCGT	GGACGATATC	CACAGGGGCT	GGCGCACCTT	CGTGTGCGT	GTGCGGGGCC	AGGACCCGCC	2170
GCCTGAGCTG	TACTTTGTCA	AGGTGGATGT	GACGGGCGCG	TACGACACCA	TCCCCCAGGA	CAGGCTCAGC	2240
GAGGTCACTG	CCAGCATCAT	CAAAACCCAG	AACACGTACT	GCGTGCCTCG	GTATGCCGTG	GTCCAGAAAG	2310
CCGCCCATGG	GCACGTCCGC	AAGGCCCTTCA	AGAGCCACGT	CTCTACCTTG	ACAGACCTCC	AGCCGTACAT	2380
GCGACAGTTC	GTGGCTCACC	TGCAGGAGAC	CAGCCCCGCTG	AGGATGCGCG	TGCTCATCGA	GCAGAGCTCC	2450
TCCCTGAATG	AGGCCAGCAG	TGGCCTCTTC	GACGTCTTCC	TACGCTTCAT	GTGCCACCAC	CGCGTGCAGA	2520
TCAGGGGCAA	GTCTTACGTC	CAGTGCCAGG	GGATCCCGCA	GGGCTCCATC	CTCTCCACGC	TGCTCTGGAG	2590
CCTGTGCTAC	GGCGACATGG	AGAACAGGCT	GTTTGCGGGG	ATTGCGCGGG	ACGGGCTGCT	CCTGCGTTTG	2660
GTGGATGATT	TCTTGTGGT	GACACCTCAC	CTCACCCACG	CGAAAACCTT	CCTCAGGACC	CTGGTCCGAG	2730
GTGTCCCTGA	GTATGGCTGC	GTGGTGAAC	TGCGGAAGAC	AGTGGTGAAC	TTCCCTGTAG	AAGACGAGGC	2800
CCTGGGTGGC	ACGGCTTTTG	TTCAGATGCC	GGCCACGGCG	CTATTCCCTT	GGTGGGGCTT	GCTGCTGGAT	2870
ACCCGGACCC	TGGAGGTGCA	GAGCGACTAC	TCCAGCTATG	CCCGGACCTC	CATCAGAGCC	AGTCTCACCT	2940
TCAACCGCGG	CTTCAAGGCT	GGGAGGAACA	TGCGTCGCAA	ACTCTTTGGG	GTCTTGCGGC	TGAAGTGTC	3010
CAGCCTGTTT	CTGGATTGCG	AGGTGAACAG	CCTCCAGACG	GTGTGCACCA	ACATCTACAA	GATCTCTCTG	3080
CTGCAGGCGT	ACAGGTTTCA	CGCATGTGTG	CTGCAGCTCC	CATTTTCATC	GCAAGTTTGG	AAGAACCCCA	3150
CATTTTCTCT	GCGCGTCATC	TCTGACACGG	CCTCCCTCTG	CTACTCCATC	CTGAAAGCCA	AGAACGACGG	3220
GATGTGCGTG	GGGCCCCAAGG	GCGCCGCGCG	CCCTCTGCCC	TCCGAGGCGG	TGCAAGTGGT	GTGCCACCAA	3290
GCATTCTCTG	TCAAGCTGAC	TGACACCGGT	GTCACCTACG	TGCCACTCCT	GGGGTCACTC	AGGACAGCCC	3360
AGACGCAGCT	GAGTCCGAAG	CTCCCGGGGA	CGACGCTGAC	TGCCCTGGAG	GCCGCAGCCA	ACCCGGCACT	3430
GCCCTCAGAC	TTCAAGACCA	TCCTGGACTG	ATGGCCACCC	GCCCAACAGC	AGGCCGAGAG	CAGACACCAG	3500
CAGCCCTGTC	ACGCCGGGCT	CTACGTCCCA	GGGAGGGAGG	GGCGGCCAC	ACCCAGGCC	GCACCGCTGG	3570
GAGTCTGAGG	CCTGAGTGAG	TGTTTGGCCG	AGGCCCTGCAT	GTCCGGCTGA	AGGCTGAGTG	TCCGGCTGAG	3640
GCCTGAGCGA	GTGTCCAGCC	AAGGGCTGAG	TGTCCAGCAC	ACCTGCCGTC	TTCACTTCCC	CACAGGCTGG	3710
CGCTCGGCTC	CACCCAGGGG	CCAGCTTTTC	CTCACCAGGA	GCCCCGCTTC	CACTCCCCAC	ATAGGAATAG	3780
TCCATCCCCA	GATTCCGCAT	TGTTCAACCC	TGCCCCCTGC	CTCCTTTGCC	TTCCACCCCC	ACCATCCAGG	3850
TGGAGACCCT	GAGAAGGACC	CTGGGAGCTC	TGGGAATTTG	GAGTGACCAA	AGGTGTGCC	TGTACACAGG	3920
CGAGGACCCT	GCACCTGGAT	GGGGTCCCT	GTGGGTCAAA	TTGGGGGAG	GTGCTGTGGG	AGTAAATATC	3990
TGAATATATG	AGTTTTCAG	TTTTGAAAA	AAAAAAAAA	AAAAAAAAA	AA		4042

Fig. 7

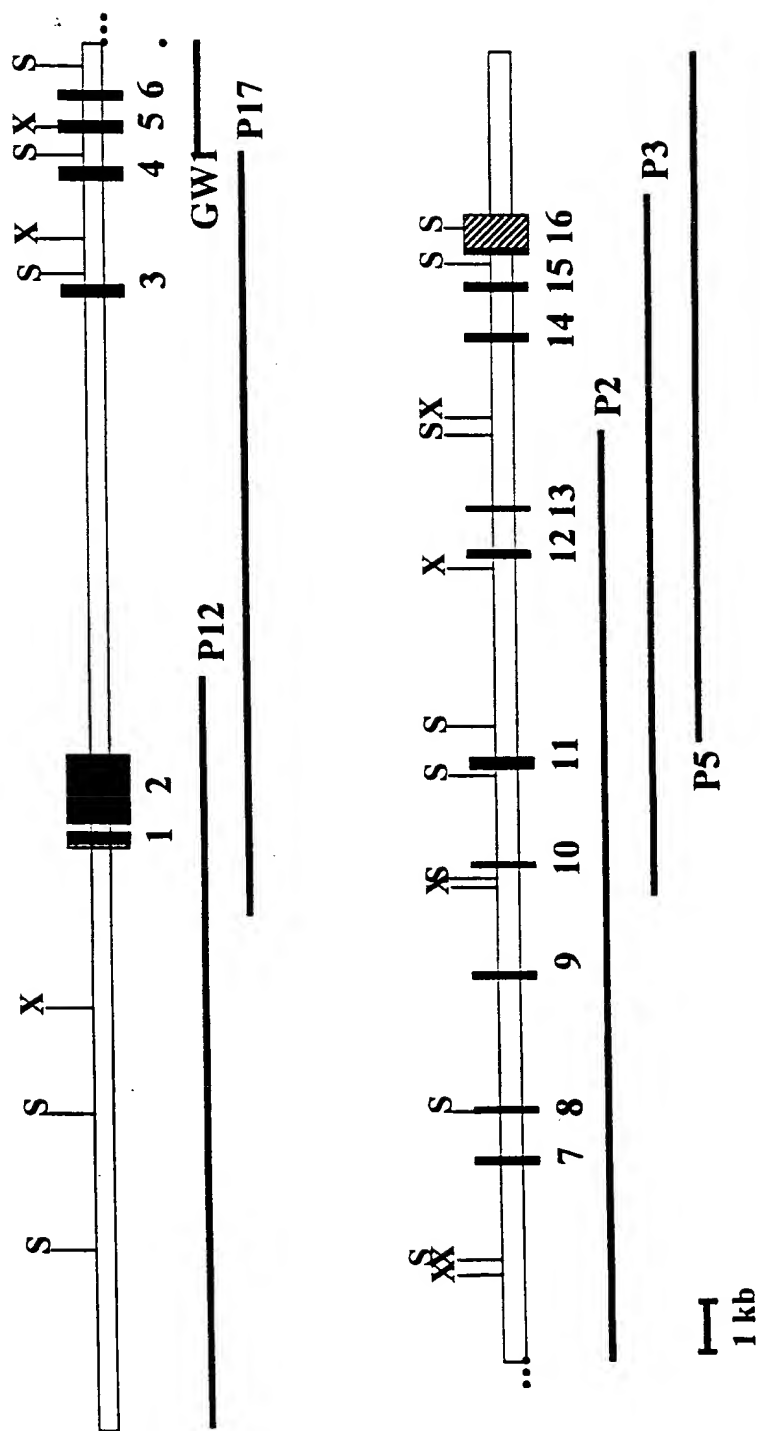


Fig. 8A

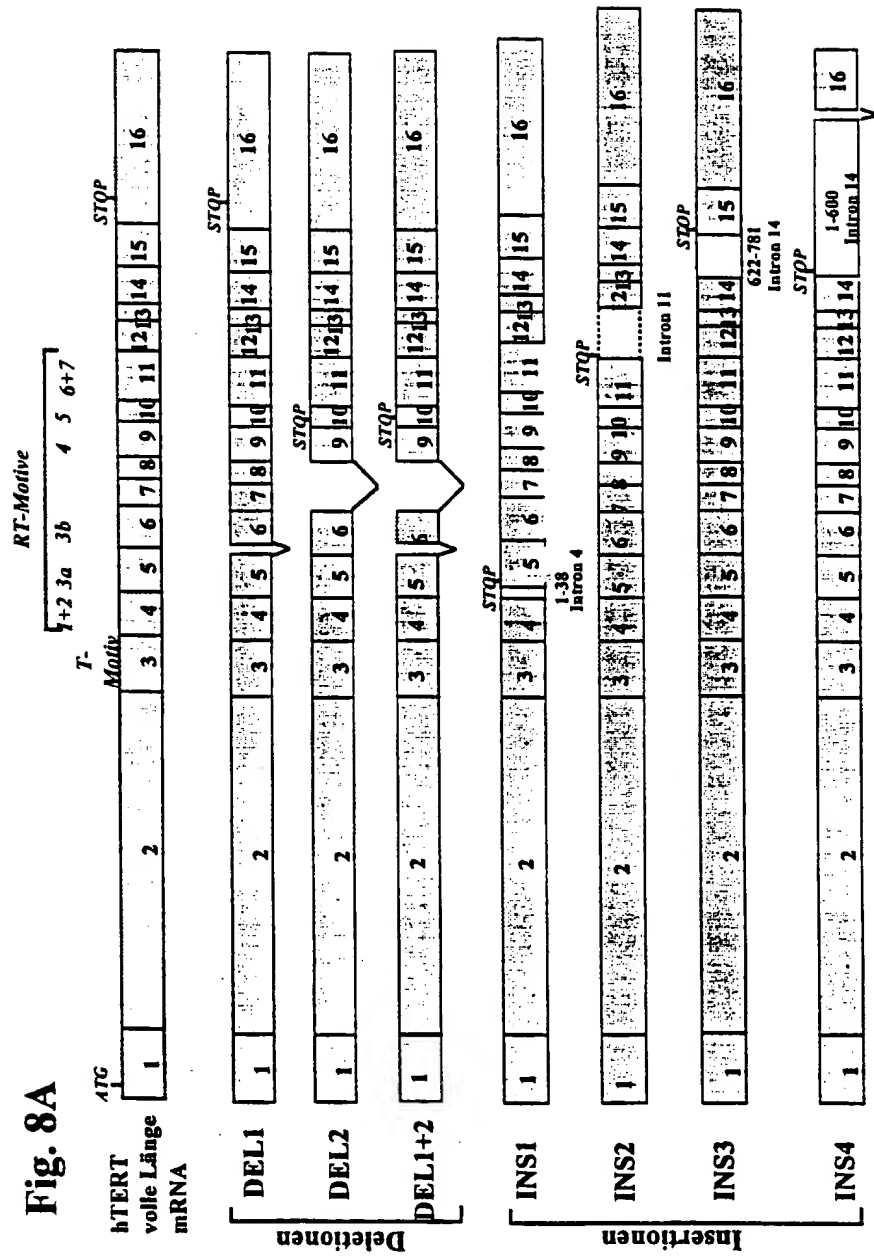
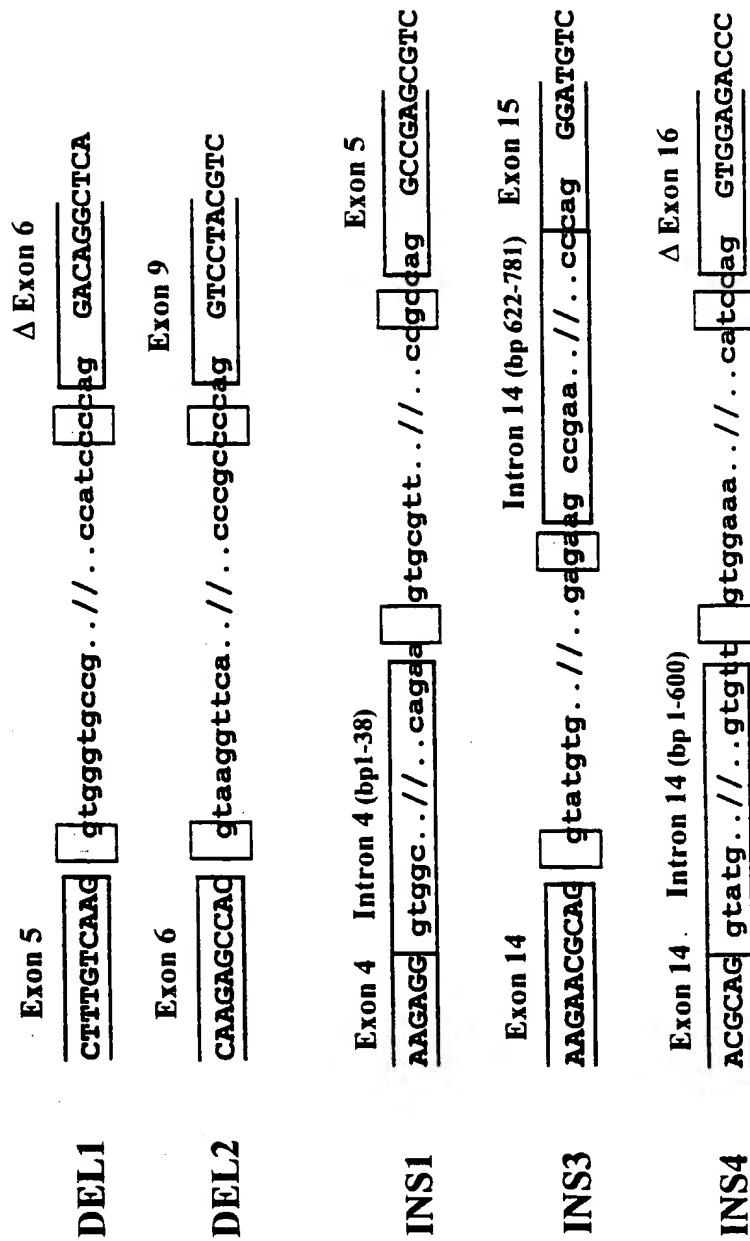


Fig. 8B



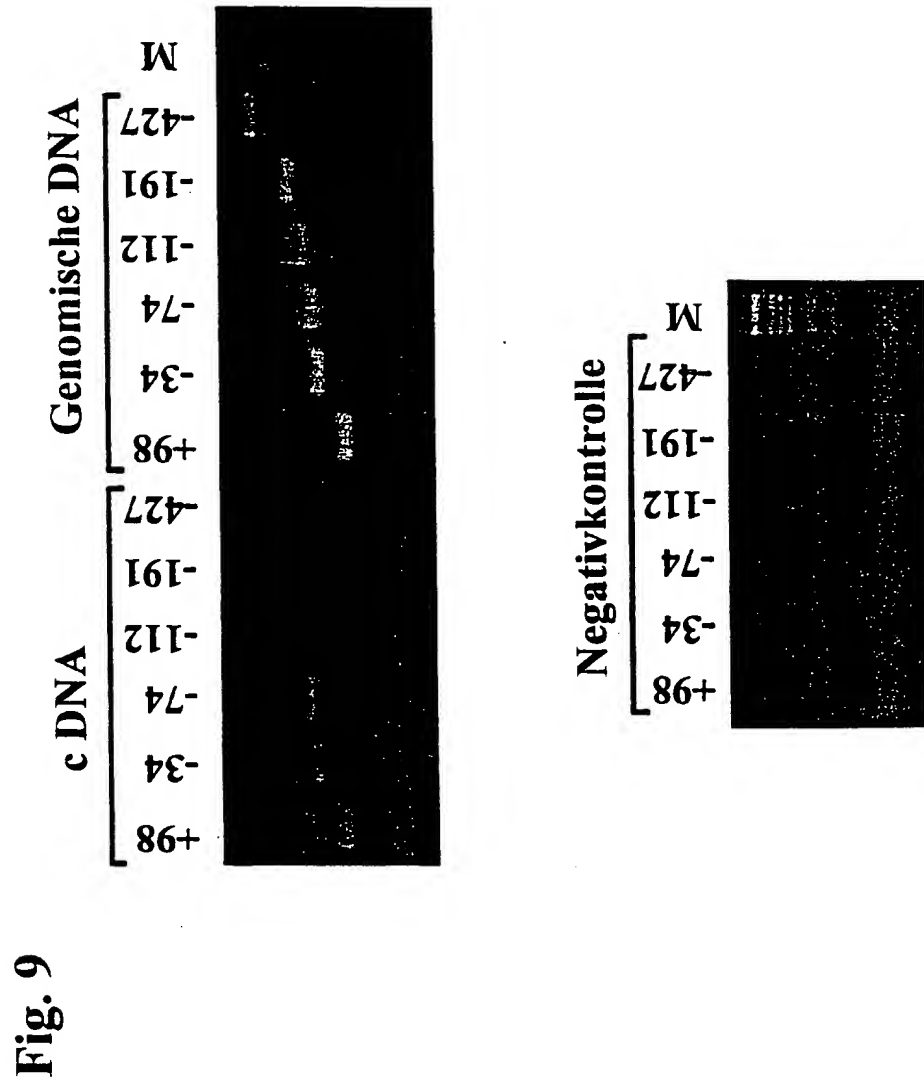


Fig. 10

ACTTGAGCCC AAGAGTTCAG GGCTACGGTG AGCCATGATT GCAACACCAC ACGCCAGCCT TGGTGACAGA -11204  
 ATGAGACCCT GTCTCAAAAA AAAAAAAAAA AATTGAAATA ATATAAAGCA TCTTCTCTGG CCACAGTGGG -11134  
 AAAAAACCAG AAATCAACAA CAAGAGGAAT TTTGAAACT ATACAAACAC ATGAAAATTA AACAAATATAC -11064  
 TTCTGAATGA CCAGTGAGTC AATGAAGAAA TTAAGAAAGG AATTGAAAAA TTTATTTAAG CAAATGATAA -10994  
 CGGAAACATA ACCTCTCAAA ACCCAGCGTA TACAGCAAAA GCAGTGCTAA GAAGGAAGTT TATAGCTATA -10924  
 AGCAGCTACA TCAAAAAAGT AGAAAAGCCA GGCGCAGTGG CTCATGCCCTG TAATCCCAGC ACTTTGGGAG -10854  
 GCCAAGGCGG GCAGATCGCC TGAGGTCAGG AGTTCGAGAC CAGCCTGACC AACACAGAGA AACCTTGTCTG -10784  
 CTACTAAAAA TACAAAATTA GCTGGGCATG GTGGCACATG CCTGTAATCC CAGCTACTCG GGAGGCTGAG -10714  
 GCAGGATAAC CGCTTGAACC CAGGAGGTGG AGGTTGCGGT GAGCCGGGAT TGCGCCATTG GACTCCAGCC -10644  
 TGGGTAACAA GAGTGAAACC CTGTCTCAAG AAAAAAAAAA AAGTAGAAAA ACTTAAAAAT ACAACCTAAT -10574  
 GATGCACCTT AAAGAAGTAG AAAAGCAAGA GCAAACTAAA CCTAAAATTG GTAAAAGAAA AGAAATAATA -10504  
 AAGATCAGAG CAGAAATAAA TGAACTGAA AGATAACAA ACRAAAGATC AACAAAATTA AAGTTGGT -10434  
 TTTTGAAGAG ATAAACAAA TTGACAACCT TTTGCCAGA CTAAGAAAAA AGGAAAGAG ACCTAAATAA -10364  
 ATAAAGTCAG AGATGAAAAA AGAGACATTA CAACTGTATC CACAGAAATT CAAAGGATCA CTAGAGGCTA -10294  
 CTATGAGCAA CTGTACACTA ATAAATTGAA AAACCTAGAA AAAATAGATA AATTCCTAGA TGCATACAC -10224  
 CTACCAAGAT TGAACCATGA AGAAATCCAA AGCCCAACA GACCAATAAC AATAATGGGA TTAAGCCAT -10154  
 AATAAAAGT CTCCTAGCAA AGAGAAGCCC AGGACCCAAT GGCTTCCCTG CTGGATTTTA CCAATCTATT -10084  
 AAAGAAGAA GAATTCCAAT CTTACTCAA CTATTCTGAA AAATAGAGGA AAGAATACTT CCAAACTCAT -10014  
 TCTACATGGC CAGTATTACC CTGATTCCAA AACCCAGCAA AAACACATCA AAAACAAAACA AACAAAAAAA -9944  
 CAGAAAGAAA GAAACTACA GGCCAATATC CCTGATGAAT ACTGATACAA AAATCCTCAA CAAACACTA -9874  
 GCAAAACAAA TTAACAACA CCTTCGAAAG ATCATTCAAT GTGATCAAGT GGGATTATT CCAGGGATGG -9804  
 AAGGATGTT CAACATATGC AAATCAATCA ATGTGATACA TCATCCCAAC AAAATGAAGT ACAAAACTA -9734  
 TATGATTATT TCACTTTATG CAGAAAAAGC ATTTGATAAA ATTCTGCACC CTTCTATGATA AAAACCTCA -9664  
 AAAAACAGG TATACAAGAA ACATACAGGC CAGGCACAGT GGCTCACACC TGCGATCCCA GCACTCTGGG -9594  
 AGGCCAAGGT GGGATGATTG CTTGGGCCCA GGAGTTTGAG ACTAGCCTGG GCAACAAAAT GAGACCTGGT -9524  
 CTACAAAAA CTTTTTAA AAATTAGCCA GGCGATGATG CATATGCCCTG TAGTCCAGC TAGTCTGGAG -9454  
 GCTGAGGTGG GAGAATCACT TAAGCCTAGG AGGTCGAGGC TGCAGTGAGC CATGAACATG TCACTGTACT -9384  
 CCAGCCTAGA CAACAGAACA AGACCCCACT GAATAAGAAG AAGGAGAAGG AGAAGGGAGA AGGGAGGGAG -9314  
 AAGGGAGGAG GAGGAGAAGG AGGAGGTGGA GGAGAAGTGG AAGGGGAAGG GGAAGGGAAA GAGGAAGAAG -9244  
 AAGAAACATA TTTCAACATA ATAAAGCCC TATATGACAG ACCGAGGTAG TATTATGAGG AAAAAGCTGAA -9174  
 AGCCTTTCTCT TAAGATCTG GAAATGACA AGGGCCCACT TTCACCACTG TGATTCAACA TAGTACTAGA -9104  
 AGTCCTAGCT AGAGCAATCA GATAAGAGAA AGAAATAAAA GGCATCCAAA CTGGAAGGA AGAAGTCAA -9034  
 TTATCCTGTT TGCAGATGAT ATGATCTTAT ATCTGAAAA GACTTAAGAC ACCACTAAAA AACTATTAGA -8964  
 GCTGAAATTT GGTACAGCAG GATACAAAA CAATGTACAA AAATCAGTAG TATTTCTATA TTCCAACAGC -8894  
 AAACATCTG AAAAAAGAAC CAAAAAGCA GCTACAAATA AAATTAACA GCTAGGAATT AACCAAAGAA -8824  
 GTGAAGATC TCTACATGA AACTATAAA ATGTTGATAA AAGAAATTGA AGAGGGCACA AAAAAAGAAA -8754  
 AGATATTCCA TGTTCTAGA TTGGAAGAA ATATCTGTT AAAATGTCCA TACTACCCA AGCAATTTAC -8684  
 AAATCAATG CAATCCCTAT TAAATACTA ATGACGTTCT TCACAGAAAT AGAAGAAACA ATTCTAAGAT -8614  
 TTGTACAGAA CCACAAAAGA CCCAGATAG CCAAGCTAT CCTGACCAA AAGAACAAA CTGGAAGCAT -8544  
 CACATTACCT GACTTCAAT TATACTACA AGCTATAGTA ACCCAAACTA CATGGTACTG GCATAAAAA -8474  
 AGATGAGACA TGGACAGAG GAACAGAATA GAGAATCCAG AAACAAATCC ATGCATCTAC AGTGAACCTCA -8404  
 TTTTGTACAA AGGTGCCAAG AACATACTTT GGGGAAAAGA TAATCTCTC AATAAATGGT GCTGGAGGAA -8334  
 CTGGATATCC ATATGCAAAA TAACATACT AGAATCTGT CTCTCACCAT ATACAAAAGC AAATCAAAAT -8264  
 GGATGAAAGG CTTAAATCTA AAACCTCAAA CTTTGCAACT ACTAAAAGAA AACACCGGAG AAACCTCTCA -8194  
 GGACATTGGA GTGGGCAAG ACTTCTTGAG TAATTCCTG CAGGCACAGG CAACCAAGC AAAACAGAC -8124  
 AAATGGGATC ATATCAAGT AAAAAGCTT TGCCAGCAA AGGAAACAAT CAACAAAGAG AAGAGACAAC -8054  
 CCACAGAATG GGAGAATATA TTTGCAACT ATTCTCTAA CAAGGAATTA ATAACAGTA TATATAAGGA -7984  
 GCTCAAACTA CTCTATAAGA AAAACAGCTA ATAAGCTGAT TTTCAAAAAT AAGCAAAAGA CTCGGGTAGA -7914  
 CATTTCTCAA AATAAGTCAT ACAAATGGCA AACAGGCATC TGAAAATGTG CTCAACACCA CTGATCATCA -7844  
 GAGAAATGCA AATCAAACT ACTATGAGAG ATCATCTCAT CCCAGTTAAA ATGGCTTTTA TTCAAAGAC -7774  
 AGGCAATAAC AAATGCCAGT GAGGATGTGG ATAAAAGGAA ACCCTTGGAC ACTGTTGGTG GGAATGGAAA -7704  
 TTGCTACCAC TATGGAGAAC AGTTTGAAG TTCCTCAAA AACTAAAAAT AAAGCTACCA TACAGCAATC -7634  
 CCATTGCTAG GTATATACTC CAAAAAGGG AATCAGTGTA TCAACAAGCT ATCTCCACTC CCACATTTAC -7564  
 TGCAGCACTG TTCATAGCAG CCAAGGTTTG GAAGCAACCT CAGTGTCAT CAACAGACGA ATGGAAAAAG -7494  
 AAAATGTGGT GCACATACAC AATGGAGTAC TACGCAGCCA TAAAAAGAA TGAGATCCTG TCAGTTGCAA -7424  
 CAGCATGGGG GGCAGTGGTC AGTATGTTAA GTGAAATAAG CCAGGCACAG AAAGACAAAC TTTTCATGTT -7354  
 CTCCCTTACT TGTGGGAGCA AAAATTAATA CAATTGACAT AGAAATAGAG GAGAATGGTG GTTCTAGAGG -7284  
 GGTGGGGGAC AGGGTGACTA GAGTCAACAA TAATTTATTG TATGTTTAA AATAACTAAA AGAGTATAAT -7214  
 TGGGTTGTTT GTAACACAAA GAAAGGATAA ATGCTTGAAG GTGACAGATA CCCATTTTAC CCTGATGTGA -7144  
 TTATTACACA TTGTATGCCT GTATCAAAAT ATCTCATGTA TGCTATAGAT ATAAACCTTA CTATATTTAA -7074  
 AATTAATAAT TTAATGGCCA GGCACGGTGG CTCATGTCCG TAATCCAGC ACTTTGGGAG CCGGAGGCGG -7004  
 GTGGATCACC TGAGGTGAGG AGTTTGAAC CAGTCTGGCC ACCATGATGA AACCTGTCT CTAATAAGA -6934  
 TACAAAAAT AGCCAGGCGT GGTGGCAGAT ACCTGTAGTC CCAACTACTC AGGAGGCTGA GACAGGAGAA -6864  
 TTGCTTGAAC CTGGGAGCGG GAGGTTGCAG TGAGCCGAGA TCATGCCACT GCAGTGCAGC CTGGGTGACA -6794  
 GAGCAAGACT CCATCTCAA ACAAAAACA AAAAAAGAAG ATTAATAATTG TAATTTTAT GTACCGTATA -6724  
 AATATATACT CTACTATATT AGAAGTTAAA AATTAACA ATTATAAAG GTAATTAACC ACTTAATCTA -6654  
 AATAAGAAC AATGTATGTG GGGTTCTAG CTTCTGAAGA AGTAAAGTT ATGGCCACGA TGGCAGAAAT -6584

Fig. 10

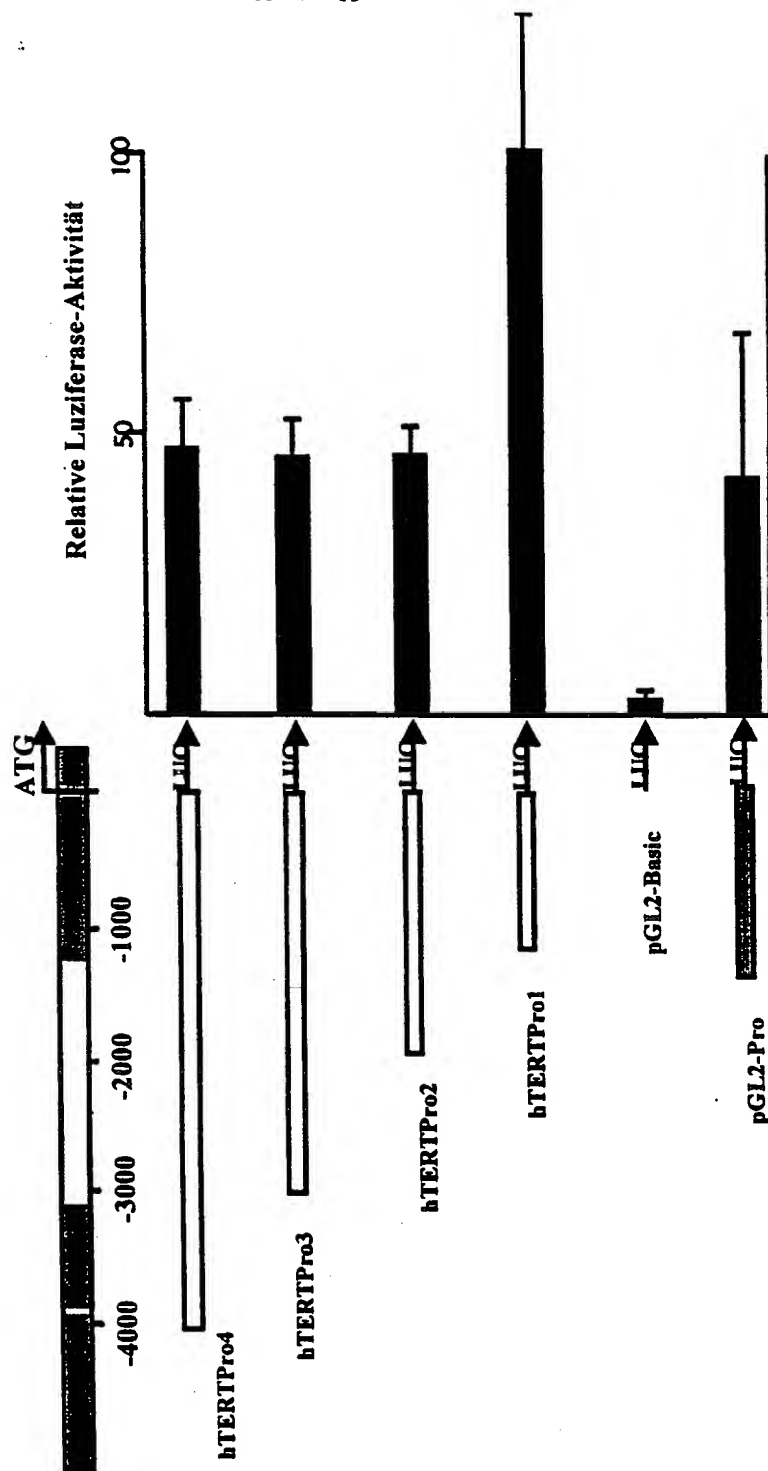
GTGAGGAGGG AACAGTGGAA GTTACTGTTG TTAGACGCTC ATACTCTCTG TAAGTGACTT AATTTTAACC -6514  
 AAAGACAGGC TGGGAGAAGT TAAAGAGGCA TTCTATAAGC CCTAAAACAA CTGCTAATAA TGGTGAAAGG -6444  
 TAATCTCTAT TAATTACCAA TAATTACAGA TATCTCTAAA ATCGAGCTGC AGAATTGGCA CGTCTGATCA -6374  
 CACCGTCTCT TCATTACAGG TGCTTTTTTT CTTGTGTGCT TGGAGATTTT CGATTGTGTG TTCGTGTTTG -6304  
 GTTAAACTTA ATCTGTATGA ATCTGAAAC GAAAAATGGT GGTGATTTCC TCCAGAAGAA TTAGAGTACC -6234  
 TGGCAGGAAG CAGGTGGCTC TGTGGACCTG AGCCACTTCA ATCTTCAAGG GTCTCTGGCC AAGACCCAGG -6164  
 TGCAAGGCAG AGGCCTGATG ACCCGAGGAC AGGAAAGCTC GGATGGGAAG GGGCGATGAG AAGCCTGCCT -6094  
 CGTTGGTGAG CAGCGCATGA AGTGCCCTTA TTTACGCTTT GCAAAGATTG CTCTGGGATC CATCTGGAAA -6024  
 AGGCGGCCAG CGGGAATGCA AGGAGTCAGA AGCCTCCTGC TCAAACCCAG GCCAGCAGCT ATGGCCGCCA -5954  
 CCCGGGCGTG TGCCAGAGGG AGAGGAGTCA AGGCACCTCG AAGTATGGCT TAAATCTTTT TTTCACCTGA -5884  
 AGCAGTGACC AAGGTGTATT CTGAGGGAAG CTTGAGTTAG GTGCCTTCTT TAAACAGAA AGTCATGGAA -5814  
 GCACCCTTCT CAAGGGAAAA CCAGACGCCC GCTCTGCGGT CATTACCTC TTTCCTCTCT CCCTCTCTTG -5744  
 CCCTCGCGGT TTCTGATCGG GACAGAGTGA CCCCGTGGGA GCTTCTCCGA GCCCGTGTCT AGGACCTCT -5674  
 TGCAAAGGGC TCCACAGACC CCCGCCCTGG AGAGAGGAGT CTGAGCCTGG CTTAATAACA AACTGGGATG -5604  
 TGGCTGGGGG CGGACAGCGA CGGCGGGATT CAAAGACTTA ATTCATGAG TAAATCAAC CTTTCCACAT -5534  
 CCGAATGGAT TTGGAATTTA TCTTAATATT TTCTTAAATT TCATCAAATA ACATTGAGG CTCGCAAAAT -5464  
 CCAAAGGCGT AAAACAGGAA CTGAGCTATG TTTGCAAGG TCCAAAGGACT TAATAACCAT GTTCAGAGGG -5394  
 ATTTTTCGCC CTAAGTACTT TTTATTGGTT TTCATAAGGT GGCTTAGGGT GCAAGGGAAA GTACACGAGG -5324  
 AGAGGCTGG GCGGCAGGGC TATGAGCAGC GCAGGGCCAC CGGGGAGAGA GTCCCCGGCC TGGGAGGCTG -5254  
 ACAGCAGGAC CACTGACCGT CCTCCCTGGG AGCTGGCCACA TTGGGCAACG CGAAGGCGCC CACGCTGCGT -5184  
 GTGACTCAGG ACCCCATACC GGCTTCCTGG GCCCACCAC ACTAACCCAG GAAGTCACGG AGCTCTGAAC -5114  
 CCGTGGAAC GAACATGACC CTTGCCCTGC TGCTTCCCTG GGTGGGTCAA GGGTAATGAA GTGGTGTGCA -5044  
 GGAATGGGCC ATGTAAATTA CACGACTCTG CTGATGGGGA CGTTCCTTC CATCATTATT CATCTCACC -4974  
 CCAAGGACT GAATGATTCC AGCAACTTCT TCGGGTGTGA CAAGCCATGA CAAAACTCAG TACAAACACC -4904  
 ACTCTTTTAC TAGGCCACCA GAGCAGGSC CACACCCCTG ATATATTAAG AGTCCAGGAG AGATGAGGCT -4834  
 GCTTTCAGCC ACCAGGCTGG GGTGACAACA GCGGCTGAAC AGTCTGTTCC TCTAGACTAG TAGACCCTGG -4764  
 CAGGCACTCC CCCAGATTCT AGGGCCTGGT TGCTGCTTCC CGAGGGCGCC ATCTGCCCTG GAGACTCAGC -4694  
 CTGGGGTGCC AACTGAGGC CAGCCCTGTC TCCACACCCT CCGCTCCAG GCCTCAGCTT CTCCAGCAGC -4624  
 TTCCTAAACC CTGGGTGGGC CGTGTTCAG CGCTACTGTC TCACCTGTCC CACTGTGTCT TGTCTCAGCG -4554  
 ACGTAGCTCG CACGGTTCCT CCTCACATGG GGTGTCTGTC TCCTTCCCA ACACCTACAT GCGTTGAAGG -4484  
 GAGGAGATTG TGCGCTCCG AGACTGGCTC CTCTGAGCCT GAACCTGGCT CGTGGCCCCG GATGCAGGTT -4414  
 CCTGGCGTCC GGCTGCACGC TGACCTCCAT TTCCAGGCGC TCCCGCTCTC CTGTCTCTG CCGGGGCGCT -4344  
 CCGGTGTGTT CTTCTGTTTC TGTGCTCCTT TCCACGTCCA GCTGCGTGTG TCTCTGCCCG CTAGGGTCTC -4274  
 GGGGTTTTTA TAGGCATAGG ACGGGGCGGT GGTGGGCCAG GCGCTCTTGG GGAATGCAA CATTGGGGT -4204  
 TGAAGTAGG AGTGCTGTC CTCACCTAGG TCCACGGGCA CAGGCCTGGG GATGGAGCCC CCGCCAGGGA -4134  
 CCCGCCCTTC TCTGCCAGC ACTTTCCTGC CCCCCTCCCT CTGGAACACA GAGTGGCAGT TTCCACAAGC -4064  
 ACTAAGCATC CTCTCCCAA AAGACCCAGC ATTGGACACC CTGGACATTT GCCCCACAGC CTTGGGAATT -3994

c-Myc  
 CACGTGACTA CGCACATCAT GTACACACTC CCGTCCACGA CCGACCCCGG CTGTTTTATT TTAATAGCTA -3924  
 CAAAGCAGGG AAATCCCTGC TAAATGTCC TTTAACAAAC AACGGGTCCA TCCGCACGGT -3854  
 GGACAGTTCC TCACAGTGAA GAGGAACATG CCGTTTATAA AGCCTGCAGG CATCTCAAGG GAATTACGCT -3784  
 GAGTCAAAC TGCCACCTCC ATGGGATACG TACGCAACAT GCTCAAAGG AAAGAATTTC ACCCCATGGC -3714  
 AGGGAGTGG TTAGGGGGGT TAAGGACGGT GGGGGCGGCA GCTGGGGGCT ACTGCACGCA CCTTTTACTA -3644  
 AAGCCAGTTT CTTGGTCTG ATGGTATTGG CTCAGTTATG GGAGACTAAC CATAGGGGAG TGGGGATGGG -3574  
 GGAACCCGGA GGCTGTGCCA TCTTTGCCAT GCCCGAGTGT CCTGGGCAGG ATAATGCTCT AGAGATGCCC -3504  
 ACGTCTGAT TCCCCAAAC CTGTGGACAG AACCCGCGCG GCCCCAGGGC CTTTGCAGGT GTGATCTCCG -3434  
 TGAGGACCCT GAGGTCTGGG ATCCTTCGGG ACTACCTGCA GGCCCGAAAA GTAATCCAGG GGTTCGGA -3364  
 AGAGCGGGC AGGAGGCTCA GAGGGGGGCA GCCTCAGGAC GATGGAGGCA GTCAGTCTGA GGCTGAAAAG -3294  
 GGAGGGAGGG CCTCGAGCCC AGGCCCTGCAA GCGCCTCCAG AAGCTGGAAA AAGCGGGGAA GGGACCTCC -3224  
 ACGGAGCCTG CAGCAGGAAG GCACGGCTGG CCCTTAGCCC ACCAGGGCCC ATCGTGGACC TCCGGCTCC -3154  
 GTGCCATAGG AGGGCACTCG CGCTGCCCTT CTAGCATGAA GTGTGTGGGG ATTTGCAGAA GCAACAGGAA -3084  
 ACCCATGCAC TGTGAATCTA GGATTATTTC AAAACAAAGG TTTACAGAAA CATCCAAGGA CAGGGCTGAA -3014  
 GTGCCTCCGG GCAAGGGCAG GGCAGGCACG AGTGATTTTA TTTAGCTATT TTATTTTATT TACTTACTTT -2944  
 CTGAGACAGA GTTATGCTCT TGTGCCCCAG GCTGGAGTGC AGCGGCATGA TCTTGGCTCA CTGCAACCTC -2874  
 CGTCTCCTGG GTTCAAGCAA TTCTCGTGCC TCAGCCTCCC AAGTAGCTGG GATTTCAGGC GTGCACCACC -2804  
 ACACCCGGCT AATTTTGTAT TTTTAGTAGA GATGGGCTTT CACCATGTTG GTCAAGCTGA TCTCAAAATC -2734  
 CTGACCTCAG GTGATCCGCC CACCTCAGCC TCCCAAAGTG CTGGGATTAC AGGCATGAGC CACTGCACCT -2664  
 GGCCTATTTA ACCATTTTAA AACTTCCCTG GGCTCAAGTC ACACCCACTG GTAAGGAGTT CATGGAGTTC -2594  
 AATTTCCCTT TTAATCAGGA GTTACCTCC TTTGATATT TCTGTAATTC TTCGTAGACT GGGGATACAC -2524  
 CGTCTCTTGA CATATTCACA GTTTCGTGTA CCACCTGTTA TCCCATGGGA CCCACTGACG GGGCAGCTGG -2454  
 GAGGCTGCAG GCTTCAGGTC CCAGTGGGGT TGCCATCTGC CAGTAGAAGC CTGATGTAGA ATCAGGGCGC -2384  
 AAGTGTGGAC ACTGTCCTGA ATCTCAATGT CTCAGTGTGT GCTGAAACAT GTAGAAATTA AAGTCCATCC -2314  
 CTCTACTCT ACTGGGATTG AGCCCCCTTC CTATCCCCCT CAGGGGGCAG AGGAGTTCCT CTCACTCCTG -2244  
 TGGAGGAAGG AATGATACTT TGTATTTTT CACTGCTGGT ACTGAATCCA CTGTTTCTG TGTGGGTTG -2174  
 TTTGTTTTGT TTTAGAGGC GGTTCCTCTC TTGTGTCTCA GGCTGGAGGG AGTGCAATGG CGCGATCTTG -2104  
 GCTTACTGCA GCCTCTGCCT CCCAGGTTCA AGTGATTCTC CTGCTCCGC CTCCCATTTG GCTGGGATTA -2034  
 CAGGACCCCG CCACCATGCC CAGCTAATTT TTTGTATTTT TAGTAGAGAC GGGGGTGGT GGGGTTCCAC -1964

ATGTTGGCCA GGCTGGTCTC GAACTTCTGA CCTCAGATGA TCCACCTGCC TCTGCCTCTT AAAGTGCTGG -1894  
GATTACAGGT GTGAGCCACC ATGCCCAGCT CAGAATTAC TCTGTTTAGA AACATCTGGG TCTGAGGTAG -1824  
GAAGCTCACC CCACTCAAGT GTTGTGGTGT TTTAACAAT-BoxGGCCAA TGGATAGAATT TTTTATTGT TGTTAGAACA -1754  
CTCTTGATGT TTTACACTGT GATGACTAAG ACATCATCAG CTTTTCAAAG ACACACTAAC TGCACCCATA -1684  
ATACTGGGGT GTCTTCTGGG TATCAGCAAT CTTCAATTGA TGCCGGGAGG CGTTTCCTCG CCATGCACAT -1614  
GGTGTTAAAT ACTCCAGCAT AATCTTCTGC TTCCATTCT TCTCTTCCCT CTTTTAAAAA TGTGTTTTCT -1544  
ATGTTGGCTT CTCTGCAGAG AACCAGTGTA AGCTACAAC TAACTTTTGT TGGAACAAAT TTTCCAACC -1474  
Sp1  
GGCCCTTTGC CCTAGTGGCA GAGACAATTC ACAAACACAG CCCTTTAAAA AGGCTTAGGG ATCACTAAGG -1404  
GGATTCTTAG AAGAGCGACC TGTAATCCTA AGTATTTACA AGACGAGGCT AACCTCCAGC GAGCGTGACA -1334  
GCCCAGGGAG GGTGCGAGGC CTGTTCAAAT GCTAGCTCCA TAAATAAAGC AATTTCCTCC GGCAGTTTCT -1264  
GAAAGTAGGA AAGGTTACAT TTAAGGTTGC GTTTGTTAGC ATTTCAAGTGT TTGCCGACCT CAGCTACAGC -1194  
ATCCCTGCAA GGCCTCGGGA GACCCAGAAG TTTCTCGCCC CTTAGATCC AAACCTTGAGC AACCCGGAGT -1124  
CTGGATTCTT GGGAAGTCCT CAGCTGTCCT GCGGTTGTGC CGGGGCCCA GGTCTGGAGG GGACCAGTGG -1054  
CCGTGTGGCT TCTACTGCTG GGCTGGAAGT CGGGCCTCAAT AGCTCTGCAG TCCGAGGCTT GGAGCCAGGT -984  
GCGTGGACCC CGAGGCTGCC CTCACCCCTG TCGGGGCGGG ATGTGACCAG ATGTTGGCCT CATCTGCCAG -914  
ACAGAGTGCC GGGGCCCAGG GTCAAGGCCG TTGTGGCTGG TGTGAGGCGC CCGGTGCGCG GCCAGCAGGA -844  
GCGCCTGGCT CCATTCCAC-BoxCCCTTTCTCG ACGGGACSp1CGC CCCGGTGGGT GATTAACAGA TTTGGGGTGG -774  
TTTGCTCATG GTGGGGACCC CTCGCCGCT GAGAACCTGC AAAGAGAAAT GACGGGCTG TGTCAGGAG -704  
CCCAAGTCGC GGGGAAGTGT TGCAGGGAGG CACTCCGGGA GGTCCCGCT GCCCGTCCAG GGAGCAATG -634  
GTCCTCGGGT TCGTAP-2CCCCAG CGCGTCTAC GCGCCTCGT CCTCCCCTTC ACGTCCGGCA TTCGTGGTGC -564  
CCGGAGCCCG ACGCCCCGCG TCCGGACCTG GAGGCAGCCC TGGGTCTCCG GATCAGGCCA GCGGCCAAAG -494  
GGTCGCCGA CGCACCTGTT CCCAGGCCT CCACATCATG GCCCCCTCCT CGGGTTACCC CACAGCCTAG Sp1 -424  
GCCGATTCTA CCTCTCTCCG CTGGGGCCCT CGCTGGCGTC CCTGCACCT GGGAGCGCA GCGGCCGCG -354  
GGCGGGGAAG CGCGGCCAG ACCCCCCGGT CGGCCGGAG CAGCTGCGCT GTCGGGGCCA GGCCGGGCTC -284  
CCAGTGGATT CGCGGGCACA GACGCCCAGG ACCGCGCTCC CGACGTGCGG GAGGGACTGG GGACCCGGGC -214  
ACCGTCTCTG CCCCTTCACC TTCCAGCTCC GCCTCCTCG CGCGGACCCGCCCTGCC GACCCCTCCC -144  
GGGTCCCCGG CCCAGCCCCC TCCGGGCCCT CCCAGCCCT CCCCCTTCCTT TCCGCGGCC CGCCCTCTCC Sp1 -74  
TCGCGGCGCG AGTTTCAGGC AGCGCTGCGT CCTGCTGCG CGACGTGGGAAG CCCTGGCCCC GGCCACCCCC -4  
GCGATG

15 / 15

Fig.: 11



## SEQUENZPROTOKOLL

&lt;110&gt; Bayer AG

5 <120> Regulatorische DNA-Sequenzen aus der 5i-Region vom Gen  
der humanen katalytischen Telomerase-Untereinheit und  
deren diagnostische und therapeutische Verwendung

10 <130> LeA32805-Ausland

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 20

15 <170> PatentIn Vers. 2.0

&lt;210&gt; 1

&lt;211&gt; 5126

20 <212> DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 1

25 gagctctgaa ccgtggaaac gaacatgacc cttgcctgcc tgcctccctg ggtgggtcaa 60  
gggtaataaa gtgggtgtgca ggaaatggcc atgtaaaatta cagcactctg ctgatgggga 120  
ccgttccttc catcattatt catcttcacc cccaaggact gaatgattcc agcaacttct 180  
tcgggtgtga caagccatga caaaactcag tacaacacacc actcttttac taggcccaca 240  
gagcacgggc cacacccctg atataattaag agtccaggag agatgaggct gctttcagcc 300  
accaggctgg ggtgacaaca gcggctgaac agtctgttcc tctagactag tagaccctgg 360  
caggcaactcc cccaaattct agggcctggg tgcctgttcc cgagggcgcc atctgcctg 420  
gagactcagc ctgggggtgcc acactgaggc cagccctgtc tccacacccct ccgcctccag 480  
gcctcagctt ctccagcagc ttccctaaacc ctgggtgggc cgtgttccag cgtactgtc 540  
tcacctgtcc cactgtgtct tgtctcagcg acgtagctcg caccggttcc cctcacatgg 600  
ggtgtctgtc tccctcccca acactcacat gcgttgaagg gaggagattc tgcgcctccc 660  
35 agactggctc ctctgagcct gaacctggct cgtggccccc gatgcagggt cctggcgctc 720  
ggctgcacgc tgacctccat tccaggcgc tccccgtctc ctgtcatctg ccggggcctg 780  
ccgggtgtgt ctctgttctc tgtgtctcct tccacgtcca gctgcgtgtg tctctgccc 840  
ctagggtctc ggggttttta taggcataag acgggggcgt ggtgggccag ggcgtctctg 900  
ggaaatgcaa catttgggtg tgaaagtagg agtgcctgtc ctacactagg tccacgggca 960  
40 caggcctggg gatggagccc ccgccaggga cccgcccttc tctgccagc acttctcctg 1020  
ccccctccct ctggaacaca gagggtgagc ttccacaagc actaagcctc ctcttcccaa 1080  
aagaccagc attggcacc ctggacattt gcccacacgc cctgggaatt cactgacta 1140  
cgcacatcat gtacacactc ccgtccacga ccgacccccc ctgttttatt ttaatagcta 1200  
caaagcaggg aaatccctgc taaaatgtcc tttacaaaac tggttaaaaca aacgggtcca 1260  
45 tccgcacggg ggacagtctc tcacagtga gaggaaacat ccgtttataa agcctgcagg 1320  
catctcaagg gaattacgct gagtcaaaac tgccacctcc atgggatacg tacgcaacat 1380  
gctcaaaaag aaagaatttc accccatggc aggggagtg ttaggggggt taaggacggg 1440  
gggggcggca gctgggggct actgcacgca ccttttacta aagccagttt cctgggtctg 1500  
atgggtattg ctcatgtatg ggagactaac cataggggag tggggatggg ggaacccgga 1560  
50 ggctgtgcca tctttgccat gcccagtggt cctgggcagg ataagtctct agagatgccc 1620  
acgtcctgat tcccccaaac ctgtggacag aaccgcgccg gccccagggc ctttgaggt 1680  
gtgatctccg tgaggacctc gaggtctggg atccttcggg actacctgca ggcccgaaaa 1740  
gtaatccagg ggttctggga agaggcgggc agggagggtca gaggggggca gcctcaggac 1800  
gatggaggca gtcagtctga ggctgaaaag ggaggggagg cctcgagccc aggcctgcaa 1860  
55 gcgcctccag aagctggaaa aagcggggaa gggaccctcc acggagcctg cagcaggaa 1920  
gcacggctgg cccttagccc accaggggcc atcgtggacc tccggcctcc gtgccatagg 1980  
agggcactcg cgtgcccctt ctgacatgaa gtgtgtgggg atttgacaga gcaacaggaa 2040  
acctatgcac tgtgaatcta ggattatttc aaaacaaaag tttacagaaa catccaagga 2100  
cagggtgaaa gtgcctccgg gcaagggcag ggcaggcacg agtgatttta tttagctatt 2160  
60 ttattttatt tacttacttt ctgagacaga gttatgctct tgttgcctcg gctggagtgc 2220  
agcggcatga tcttggctca ctgcaacctc cgtctcctgg gtccaagcaa tctcgtgccc 2280  
tcagccctcc aagtagctgg gatttcaggc gtgcaccacc acaccggcct aattttgtat 2340  
ttttagtaga gatgggcttt caccatgttg gtcaagctga tctcaaaatc ctgacctcag 2400  
gtgatccgcc cactcagcc tcccaaatg ctgggattac aggcagtagc cactgcacct 2460

2 / 18

5 ggccatttta accattttta aacttccctg ggctcaagtc acaccactg gtaaggagtt 2520  
 catggagttc aatttccctt ttactcagga gttaccctcc tttgatatct tctgtaattc 2580  
 ttccgtagact ggggatacac cgtctcttga catattcaca gtttctgtga ccacctgtta 2640  
 tccccatggga cccactgcag gggcagctgg gaggtgcag gcttcaggtc ccagtggggg 2700  
 10 tggcatcrgc cagtagaaac ctgatgtaga atcagggcgc aagtgtggac actgtcctga 2760  
 atctcaatgt ctctagtgtg gctgaaacat gtagaaatta aagtcacatcc ctctactctc 2820  
 actgggattg agccccctcc ctatccccc ccaggggcag aggagttcct ctactcctg 2880  
 tggaggaaagg aatgatactt tgttattttt cactgctggt actgaatcca ctgtttcatt 2940  
 tgttggtttg tttgttttgt tttgagaggg ggtttcactc ttgttgctca ggctggaggg 3000  
 15 agtgcaatgg cgcgatcttg gcttactgca gcctctgctt cccagggttca agtgattctc 3060  
 ctgcttccgc ctccccattg gctgggatta caggcacccg ccaccatgcc cagctaattt 3120  
 tttgtatttt tagtagagac ggggttgggt ggggttcacc atgttggtcca ggctggctcc 3180  
 gaacttctga cctcagatga tccacctgcc tctgcctcct aaagtgtctg gattacaggt 3240  
 gtgagccacc atgcccagct cagaatttac tctgtttaga aacatctggg tctgaggtga 3300  
 20 tgttagaaca ctcttgatgt ttacactgt gatgactaag acatcatcag cttttcaaa 3360  
 acacactaac tgcaccata atactgggtt gtcttctggg tatcagcaat cttcattgaa 3480  
 tgcggggagg cgtttcctcg ccactgcacat ggtgttaatt actccagcat aatcttctgc 3540  
 ttccatttct tctcttccct cttttaaaat tctgttttct atgttggtct ctctgcagag 3600  
 aaccagtgtg agctacaact taacttttgt tggaaacaaat ttccaaacc gcccttttgc 3660  
 cctagtggca gagacaattc acaaacacag ccttttaaaa aggtcttagg atcactaaag 3720  
 ggatttctag aagagcgacc tgaatccta agtatttaca agacgaggct aacctccagc 3780  
 25 gagcgtgaca gccacgggag ggtgcgaggg ctgttcaaat gctagctcca taaataaaag 3840  
 aatttctctc ggcagtttct gaaagtagga aaggttacat ttaaggttgc gtttggttagc 3900  
 atttctagtgt ttgccgacct cagctacagc atccctgcaa ggcctcggga gaccagaag 3960  
 ttctcgcgcc ccttagatcc aaacttgagc aaccggagt ctggattcct gggaagtcc 4020  
 cagctgtcct gcggtgtgct cggggcccca ggtctggagg ggaccagtgg ccgtgtggct 4080  
 tctactgctg ggttggaagt cgggctcctt agctctgcag tccgaggctt ggagccaggt 4140  
 30 gcttggaacc cgaggctgcc ctccacctg tgcggggcgg atgtgaccag atgttggtct 4200  
 catctgccag acagagtgcc gggggccagg gtcaaggccg ttgtggctgg tgtgagggc 4260  
 ccggtgcgag gccagcagga gcgcctggct ccatttccca ccttttctcg acgggaccgc 4320  
 cccggtgggt gattaaacaga tttgggggtg tttgtctatg gtgggggacc ctgcgcgct 4380  
 gagaacctgc aaagagaaat gacgggctg tgtcaaggag cccaagtgc ggggaagtgt 4440  
 35 tgcaggaggg cactccggga ggtcccgctt gcccgctcag ggagcaatgc gtcctcgggt 4500  
 tctgtccccc cgcgctctac gcgcctccgt cctcccttcc acgtccggca tctgtgtgct 4560  
 ccggagcccg acgccccgct tccggacctg gaggcagccc tgggtctccg gatcaggcca 4620  
 gcgcccaaa ggtcgccgca cgcacctgtt cccagggctt ccacatcatg gccctcctct 4680  
 cgggttacc caccagctag gccgattcga cctctctccg ctggggccct cgtggcgctc 4740  
 40 cctgcacctt gggagcgaga gcggcgcgcg ggcgggggag cgcggcccag acccccggtt 4800  
 ccgccccgag cagctgcgct gtcggggcca ggcggggctc ccagtggatt cgcggggcct 4860  
 gacgcccagg accgctctcc ccacgtggcg gagggactgg ggaccggggc acccgtctct 4920  
 ccccttacc ttcagctcc gcctcctccg cgcggacccc gcccgctccc gaccctctcc 4980  
 ggggtccccg cccagcccc tccgggacct cccagcccc ccccttctct tccgctggcc 5040  
 45 cgccctctcc tgcggggcag agttttaggc agcgtctgct cctgctgcgc acgtgggaa 5100  
 ccctggcccc ggccaccccc gcgatg 5126

<210> 2  
 <211> 4042  
 <212> DNA  
 50 <213> Homo sapiens

<400> 2  
 gtttcaggca gcgctgcgct ctgctgcgca cgtgggaagc cctggccccg gccacccccg 60  
 55 cgatgcgcgc cgctccccgc tgcgagcccg tgcgctccct gctgcgcagc cactaccgcg 120  
 aggtgctgcc gctggccacg ttcgtgcggc gcctggggcc ccagggtctg cggctggtgc 180  
 agcgcgggga cccggcggtt ttccgcgcgc tgggtggcca gtgcctggtg tgcgtgccct 240  
 gggacgcacg gccgcccccc gccgccccct ccttccgcca ggtgtcctgc ctgaaggagc 300  
 tgggtggccc agtgctgcag aggtgtgtcg agcgcggcgc gaagaacgtg ctggccttcc 360  
 60 gcttcgcgct gctggacggg gcccgcgggg gccccccga ggcccttacc accagcgtgc 420  
 gcagctacct gcccaacacg gtgaccgacg cactgcgggg gagcggggcg tgggggctgc 480  
 tgctgcgcgc cgtgggggac gacgtgtgtg ttcacctgct ggcaagctgc gcgctctttg 540  
 tgcgtgtggc tcccagctgc gcctaccagg tgtgcgggccc gccgctgtac cagctcggcg 600  
 ctgccactca ggccccggcc ccgccacacg ctagtggacc ccgaaggcgt ctgggatgctg 660  
 65 aacgggcctg gaaccatagc gtcaggaggg ccgggggtccc cctgggacct ccagcccccg 720  
 gtgcgaggag gcgcgggggc agtggccagc gaagtctgct gttgcccaag aggccagggc 780

3 / 18

gtggcgctgc ccctgagccg gagcggacgc ccgttgggca ggggtcctgg gccaccccg 840  
 gcaggacgcg tggaccgagt gaccgtggtt tctgtgtggt gtcacctgcc agacccgcgc 900  
 aagaagccac ctctttggag ggtgcgctct ctggcacgcg ccactccac ccactccgtgg 960  
 gccgccagca ccacgcgggc ccccatcca catcgcgcc accacgtccc tgggacacgc 1020  
 5 ctgtgtcccc ggtgtacgcc gagaccaagc acttcctcta ctctcaggc gacaaggagc 1080  
 agctgcggcc ctcttctcta ctacgtcttc tgaggcccag cctgactggc gctcggaggc 1140  
 tcgtggagac catctttctg ggttccaggc cctggatgcc agggactccc cgcagggtgc 1200  
 ccgcctgcc ccagcgctac tggcaaatgc ggccctgtt tctggagctg cttgggaacc 1260  
 acgcgcagtg ccctacggg gtgctcctca agacgcactg cccgctgcga gctgcggtca 1320  
 10 ccccgagcgc cgggtgtctgt gcccgggaga agccccaggg ctctgtggcg gcccccagg 1380  
 aggaggacac agaccccggt cgctgggtgc agctgctccg ccagcacagc agccccctggc 1440  
 aggtgtacgg ctctgtggg gcctgcctgc gccggctggt gccccaggc ctctggggct 1500  
 ccaggcacia cgaacgcgc ttctcaggga acaccaagaa gttcatctcc ctggggaagc 1560  
 atgccaaagt ctctgtcag gagctgagc ggaagatgag cgtgcgggac tgcgcttggc 1620  
 15 tgcgagggag cccagggggt ggctgtgttc cggccgcaga gcacgtctg cgtgaggaga 1680  
 tcttgcccaa gttcctgcac tggctgatga gtgtgtacgt cgtcagagtg ctacgggtctt 1740  
 tcttttatgt cacggagacc acgtttcaaa agaacaggct ctttttctac cgggaagagt 1800  
 tctggagcaa gttgcaaaagc attggaatca gacagcact gaagagggtg cagctgcggg 1860  
 agctgtcgga agcagaggtc aggcagcatc gggaagccag gcccgccctg ctgacgtcca 1920  
 20 gactccgctt catccccaa cctgacgggc tgcgcccgat tgtgaacatg gactacgtcg 1980  
 tgggagccag aacgtttccg agagaaaaga gggccgagcg tctcacctcg aggggtgaag 2040  
 cactgttcag cgtgtcaac tacgagcggg cgcgcccgc cggcctcctg ggccctctg 2100  
 tgctgggctt ggacgatac cacagggcct ggccgacctt cgtgtgtcgt gtgcgggccc 2160  
 aggacccgcc gcctgagctg tactttgtca aggtggatgt gacggggcg taccagacca 2220  
 25 tccccagga caggctcacg gaggtcatc ccagcatcat caaacccag aacacgtact 2280  
 gcgtgcgtcg gtatgccgtg gtccagaagg ccgccatgg gcacgtccg aaggccttca 2340  
 agagccacgt ctctacctt acagacctcc agccgtacat gcgacagttc gtggctcacc 2400  
 tgcaggagac cagcccgctg agggatgccc tcgtcatcga gcagagctcc tccctgaatg 2460  
 aggccagcag tggcctcttc gacgtcttcc tacgcttcat gtgccaccac gccgtgcgca 2520  
 30 tcaggggcaa gtctacgtc cagtgcagg ggatcccga gggctccat ctctccacgc 2580  
 tgctctgtag cctgtgtac ggcgacatg agaacaagct gtttgcgggg attcggcggg 2640  
 acggggtgct cctgcgtttg gtggatgatt tcttgttggg gacacctcac ctacccacg 2700  
 cgaaaacctt cctcaggacc ctggtccgag gtgtccctga gtatggctgc gtgttgaact 2760  
 35 tgcggaagac agtgggtgaac ttccctgtag aagacgaggc cctgggtggc acggcttttg 2820  
 ttcagatgcc ggcccacggc ctattcccct ggtgcccgt gctgttggt acccgagccc 2880  
 tggaggtgca gacgactac tccagctatg cccggacctc catcagagcc agtctcact 2940  
 tcaaccgcgg cttaaggct gggaggaaca tgcgtcgcaa actctttggg gtcttgcggc 3000  
 tgaagtgtca cagcctgtt ctggatttgc aggtgaacag cctccagacg gtgtgcacca 3060  
 40 acatctacaa gatcctcctg ctgcaggcgt acaggtttca cgcagtgtg ctgcagctcc 3120  
 catctcatca gcaagtttg aagaacccca catcttctct gcgctcatc tctgacagcg 3180  
 cctccctctg ctactccatc ctgaaagcca agaacgcagg gatgtcgtg ggggccaagg 3240  
 gcgcccgcgg cctctgtccc tccgaggcgg tgcagtggct gtgccaccaa gcatctctgc 3300  
 tcaagctgac tcgacacccc gtccactcct ggggtcactc aggcacagcc 3360  
 45 agacgcagct gagtcggaag ctcccgggga cgcgctgac tgcctggag gccgcagcca 3420  
 acccggcact gccctcagac ttcaagacca tccctgactg atggccaccc gccacagcc 3480  
 agggcgagag cagacaccag cagccctgtc acgcccggct ctacgtccc gggaggagg 3540  
 ggccggccac acccaggccc gcaccgctgg gactctgagg cctgagttag tgtttggccg 3600  
 aggcctgcat gtccggctga aggtgtagt tccggctgag gcctgagcga gtgtccagcc 3660  
 aagggtgtag tgtccagcac acctgcccgt ttacttccc cacaggctgg cgtcgggtc 3720  
 50 caccacaggg ccagcttttc ctacccagga gcccggttcc cactccccac ataggaatag 3780  
 tccatccca gattcgccat tgttcacccc tcgcccctgc ctcttttggc tccaccccc 3840  
 accatccagg tggagacctt gagaaggacc ctgggagctc tgggaatttg gactgaccaa 3900  
 aggtgtgccc tgtacacagg cagggacctt gcacctggat ggggttccct gtgggtcaaa 3960  
 ttggggggag gtgctgtggg agtaaaatc tgaatatatg agtttttcag ttttgaaaaa 4020  
 55 aaaaaaaaaa aaaaaaaaaa aa 4042

<210> 3  
 <211> 11276  
 <212> DNA  
 60 <213> Homo sapiens

<400> 3  
 acctgagccc aagagttcaa ggctacgggtg agccatgatt gcaacaccac acgccagcct 60  
 tgggtgacaga atgagacctt gtctcaaaaa aaaaaaaaaa aattgaaata atataaagca 120  
 65 tcttctctgg ccacagtgga acaaaaccag aaatcaacaa caagagggaat tttgaaaact 180

	atcaaaacac	atgaaaaatta	aacaatatat	ttctgaatga	ccagtgaagtc	aatgaagaaa	240
	ttaaaaagga	aattgaaaaa	tttattttaag	caaatgataa	cggaacata	acctctcaaa	300
	acccacggta	tacagcaaaa	gcagtgtctaa	gaagggaagt	tatagctata	agcagctaca	360
	tcaaaaaagt	agaaaaagcca	ggcgcagtg	ctcatgcctg	taatcccagc	actttgggag	420
5	gccaaaggcg	gcagatcgcc	tgaggctcagg	agttcgagac	cagcctgacc	aacacagaga	480
	aaccttctcg	ctactaaaaa	tacaaaatta	gctgggcatg	gtggcacatg	cctgtaatcc	540
	cagctactcg	ggaggctgag	gcaggataac	cgcttgaacc	caggaggtgg	aggttgcggt	600
	gagccgggat	tgccgcattg	gactccagcc	tgggtaaaca	gagtgaacc	ctgtctcaag	660
	aaaaaaaaaa	aagtagaaaa	acttaaaaat	acaacctaat	gatgcacctt	aaagaactag	720
10	aaaagcaaga	gcaaaactaaa	cctaaaattg	gtaaaagaaa	agaaataata	aagatcagag	780
	cagaaaataa	tgaaactgaa	agataacaat	acaaaagatc	aacaaaatta	aaagtgtggt	840
	ttttgaaaag	ataaacaaaa	ttgacaaacc	tttgcccaga	ctaagaaaaa	aggaaagaa	900
	acctaaataa	ataaagtcag	agatgaaaaa	agagacatta	caactgatac	cacagaaatt	960
	caaaggatca	ctagaggcta	ctatgagcaa	ctgtactata	ataaattgaa	aaacctagaa	1020
15	aaaatagata	aattcctaga	tgcatacaac	ctaccaagat	tgaacctatg	agaaatccaa	1080
	agcccaaaaca	gaccaataac	aataatggga	ttaaagccat	aataaaaaagt	ctcctagcaa	1140
	agagaagccc	aggacccaat	ggcttccctg	ctggatttta	ccaatcatct	aaagaagaa	1200
	gaattccaat	cctactcaaa	ctattctgaa	aaatagagga	aagaataact	ccaaactcat	1260
	tctacatggc	cagttattacc	ctgattccaa	aaccagacaa	aaacacatca	aaaacaaaca	1320
20	aacaaaaaaa	cagaaagaaa	gaaaactaca	ggccaatatc	cctgatgaat	actgatacaa	1380
	aaatcctcaa	caaaacacta	gcaaaccaaa	ttaaacacaa	ccttcgaaa	atcattcatt	1440
	gtgatcaagt	gggatttatt	ccagggtatg	aaggatggtt	caacatatgc	aaatcaatca	1500
	atgtgatata	tcattcccaac	aaaatgaagt	acaaaaacta	tatgattatt	tcactttatg	1560
	cagaaaaagc	atttgatata	attctgcacc	cttcatgata	aaaaccctca	aaaaaccagg	1620
25	tatacaagaa	acatacaggc	caggcacagt	ggctcacacc	tgcatccca	gcactctggg	1680
	aggccaaggt	gggatgatgt	cttgggcca	ggagtgtgag	actagcctgg	gcaacaaaat	1740
	gagacctggt	ctacaaaaaa	cttttttaaa	aaattagcca	ggcatgatgg	catatgcctg	1800
	tagtcccagc	tagtctggag	gctgaggtgg	gagaatcact	taagcctagg	aggtcgaggc	1860
	tgcatgagc	catgaacatg	tcactgtact	ccagcctaga	caacagaaca	agaccccatc	1920
30	gaataagaag	aaggagaagg	agaagggaga	agggaggagg	aaggaggagg	gaggagaagg	1980
	aggaggtgga	ggagaagtgg	aaggggaagg	ggaaaggaaa	gaggagaagg	aagaacata	2040
	tttcaacata	ataaaagccc	tatatgacag	accgaggtag	tattatgagg	aaaaactgaa	2100
	agccttctct	ctaatgctg	gaaaatgaca	agggcccact	ttcaccactg	tgattcaaca	2160
	tagtactaga	agtccctagc	agagcaatca	gataagagaa	agaaataaaa	ggcatccaaa	2220
35	ctggaaagga	agaagtcaaa	ttatctctgt	tgcatgatgt	atgatcttat	atctggaaaa	2280
	gacttaagac	accactaaaa	aactattaga	gctgaatttt	ggtacagcag	gatacaaaa	2340
	caatgtacaa	aaatcagtag	tatttctata	ttccaacagc	aaacaatctg	aaaaagaaac	2400
	caaaaaagca	gctacaaaata	aaattaaaca	gctaggaatt	aacaaaagaa	gtgaagatc	2460
	tctacaatga	aaactataaa	atgttgataa	aagaaattga	agagggcaca	aaaaagaaa	2520
40	agatattcca	tgttcataga	ttggaagaat	aaatactgtt	aaaatgtcca	tactacccaa	2580
	agcaatttac	aaattcaatg	caatccctat	taaaatacta	atgacgttct	tcacagaaat	2640
	agaagaaaca	attctaatag	ttgtacagaa	ccacaaaaga	cccagaatag	ccaaagctat	2700
	cctgaccaaa	aagaacaaaa	ctggaagcat	cacattacct	gacttcaaat	tatactacaa	2760
	agctatagta	acccaaacta	catgggtactg	gcataaaaaac	agatgagaca	tggaaccagag	2820
45	gaacagaata	gagaatccag	aaacaaatcc	atgcatctac	agtgaactca	tttttgacaa	2880
	aggtgccaa	aacatacttt	ggggaaaaaga	taatctcttc	aataaatggt	gctggaggaa	2940
	ctgggatatc	atatgcaaaa	taacaatact	agaactctgt	ctctcaccat	atacaaaagc	3000
	aaatcaaaat	ggatgaaagg	cttaaatcta	aaacctcaaa	ctttgcaact	actaaaagaa	3060
	aacaccggag	aaactctcca	ggacattgga	gtgggcaaa	acttcttgag	taattccctg	3120
50	caggcacagg	caaccaaaagc	aaaaacagac	aaatgggac	atatcaagtt	aaaaagcttc	3180
	tgcccagcaa	aggaaacaat	caacaaagag	aagagacaac	ccacagaatg	ggagaatata	3240
	tttgcaaaact	attcatctaa	caaggaatta	ataaccagta	tataataagg	gctcaaaacta	3300
	ctctataaga	aaaacacctta	ataagctgat	tttcaaaaaa	aagcaaaaaga	tctgggtaga	3360
	catttctcaa	aataagtcata	acaaatggca	aacaggcatc	tgaaaatgtg	ctcaacacca	3420
55	ctgatcatca	gagaaatgca	aatcaaaact	actatgagag	atcatctcat	cccagttaaa	3480
	atgggtttta	ttcaaaaagac	aggcaataac	aaatgccagt	gaggatgtgg	ataaaaggaa	3540
	acccttgagc	actgttgggtg	ggaatggaaa	ttgctaccac	tatggagaac	agtttgaag	3600
	ttctcaaaa	aactaaaaat	aaagctacca	tacagcaatc	ccattgctag	gtatatactc	3660
	caaaaaagg	aatcagtgta	tcaacaagct	atctccactc	ccacatttac	tgacgactg	3720
60	ttcatagcag	ccaaggtttg	gaagcaacct	cagtgctcat	caacagacga	atggaaaaag	3780
	aaaatgtggt	gcacatacac	aatggagtac	tacgcagcca	taaaaaagaa	tgagatcctg	3840
	tcagtgtcaa	cagcatgggg	ggcactgggtc	agtatgttaa	gtgaaataag	ccaggcacag	3900
	aaagacaaac	ttttcatggt	ctcccttact	tgtgggagca	aaaattaaaa	caattgacat	3960
	agaaatagag	gagaatgggtg	gttctagagg	gggtgggggac	aggggtgacta	gagtcacaaa	4020
65	taattttatg	tatgttttaa	aataactaaa	agagtataat	tgggttggtt	gtaacacaaa	4080

5 gaaaggataa atgcttgaag gtgacagata ccccatTTac cctgatgtga ttattacaca 4140  
 ttgtatgcct gtatcaaaat atctcatgta tgctatagat ataaacccta ctatattaaa 4200  
 aattaaaaatt ttaattggcca ggcacggttg ctcatgtccg taatcccagc accttgggag 4260  
 gccgagggcg gtggatcacc tgaggtcagg agtttgaac cagtcctggc accatgatga 4320  
 aacctgtct ctactaaaga tacaataatt agccaggcgt ggtggcacat acctgtagtc 4380  
 ccaactactc agggaggctga gacaggagaa ttgcttgaac ctgggaggcg gaggttgacg 4440  
 tgagccgaga tcatgccact gcactgcagc ctgggtgaca gagcaagact ccatctcaaa 4500  
 acaaaaaaaa aaaaaagaag attaaaaatt taatttttat gtaccgtata aatataact 4560  
 ctactatatt agaagttaaa aattaaaaa attataaaag gtaattaaacc acctaatcta 4620  
 10 aaataagaac aatgtatgtg gggtttctag ctctgaaga agtaaaagt ttggccacga 4680  
 tggcagaaat gtgaggaggg aacagtggaa gttactgttg ttagacgctc atactctctg 4740  
 taagtgaatt aattttaacc aaagacaggc tgggagaagt taaagaggca ttctataagc 4800  
 cctaaaaaaa ctgctaataa tggtgaaaag taatctctat taattaccaa taattacaga 4860  
 tatctctaaa atcagagctgc agaattggca cgtctgatca caccgtcctc tcatctcacg 4920  
 15 tgcctttttt cttgtgtgtg tggagatttt cgattgtgtg ttcgtgtttg gttaaactta 4980  
 atctgtatga atcctgaac gaaaaatggt ggtgatttcc tccagaagaa ttagagtacc 5040  
 tggcaggaa gaggctgctc tgggacctg agccacttca atcttcaagg gtctctggcc 5100  
 aagaccagg tgcaaggcag aggcctgatg acccgaggac aggaaagctc ggatgggaag 5160  
 gggcgatgag aagcctgcct cgttggtgag cagcgcatga agtgccctta ttacgcttt 5220  
 20 gcaaagattg ctctggatc catctgaaa agcgggccag cgggaatgca aggagtcaga 5280  
 agcctcctgc tcaaacccag gccagcagct atgggcccc cccggcgctg tggcagaggg 5340  
 agaggagtca aggcacctcg aagtatggct taaatctttt ttccactga agcagtgacc 5400  
 aaggtgtatt ctgagggaag cttgagttag gtgccttctt taaaacagaa agtcatgaa 5460  
 gcacccttct caagggaaaa ccagacgccc gctctgctg ctttaccctc ttctctctct 5520  
 25 cctctctctg cctcgcggt tcttgatcgg gacagagtga ccccgctgga gcttctccga 5580  
 gcccgctgct aggaacctct tgcaaaaggc tccacagacc cccgctctg agagaggagt 5640  
 ctgagcctgg cttaataaca aactgggatg tggctggggg cggacagcga cggcgggatt 5700  
 caaagactta attccatgag taaattcaac ctttccacat ccgaatggat ttggatttta 5760  
 30 tcttaatat ttcttaaat tcatcaata acattcagga ctgcagaaat ccaaaggcgt 5820  
 aaaaacaggaa ctgagctatg tttgccagg tccaaggact taataacat gttcagaggg 5880  
 atttttcgcc ctaagtactt tttattggtt ttcataagg ggcctagggt gcaaggga 5940  
 gtacacgagg agaggcctgg gcggcagggc tatgagcagc gaggggccac cggggagaga 6000  
 gtccccgccc tgggaggctg acagcaggac cactgacct cctccttggg agctgccaca 6060  
 35 ttgggcaacg cgaaaggcgc cacgctgctg gtgactcagg accccatacc ggcttctctg 6120  
 gcccaccac actaacccag gaagtcacgg agctctgaac ccgtggaaac gaacatgacc 6180  
 cttgcctgccc tgcttccctg ggtgggtcaa gggtaataa gtggtgtgca ggaatggcc 6240  
 atgtaaaata cacgactctg ctgaggggga ccgttctctc catcattatt catcttcacc 6300  
 cccaaggact gaatgattcc agcaacttct tccgggtgtga caagccatga caaaactcag 6360  
 40 tacaacacc actcttttac tagggccaca gagcacggsc cacaccctg atatatag 6420  
 agtccaggag agatgaggct gctttcagcc accaggctgg ggtgacaaca gcggctgaa 6480  
 agtctgttcc tctagactag tagaccctgg caggcactcc cccagattct agggcctgg 6540  
 tgctgcttcc cgaggggccc atctgcccct gagactcagc ctggggtgccc acactgaggc 6600  
 cagccctgct tccacacct cccgctccag gcctcagctt ctccagcagc ttcttaaac 6660  
 45 ctgggtgggc cgtgttccag cgtactgtc tcacctgtcc cactgtgtct tgtctcagc 6720  
 acgtagctcg cacggttctt cctcacatgg ggtgtctgtc tcttcccca acactcacat 6780  
 gcgttgaagg gaggagattc tgcgctccc agactggctc ctctgagcct gaacctggct 6840  
 cgtggccccc gatgcagggt cctggcgtcc ggtgcacgc tgacctccat ttccaggcgc 6900  
 tcccgtctc ctgtcatctg ccggggcctg ccggtgtgtt cttctgtttc tgtgtcctt 6960  
 50 tccacgtcca gctgctgtg tctctgccc ctagggtctc ggggttttta taggcatagg 7020  
 acggggggcgt ggtgggcccag ggcgtctctg ggaatgcaa catttgggtg tgaagtagg 7080  
 agtgccctgtc ctacactagg tccacgggca caggcctggg gatggagccc ccgcccagga 7140  
 cccgcccctc tctgcccagc actttctctg cccctccct ctggaacaca gaggggcagt 7200  
 ttccacaagc actaagcctc ctcttcccaa aagaccagc attggcacc ctggacattt 7260  
 55 gcccacagc cctgggaatt cacgtgacta cgcacatcat gtacacact cctccacga 7320  
 ccgacccccg ctgttttatt ttaatagcta caaagcagg aaatccctgc taaaatgtcc 7380  
 ttaacaaaac tgggttaaaa aacgggtcca tccgcacggt ggacagtcc tcacagtga 7440  
 gaggaacatg ccgtttataa agcctgcagg catctcaagg gaattacgct gagtcaaac 7500  
 tgccacctcc atgggatacg tacgcaacat gctcaaaaag aaagaatttc accccatggc 7560  
 60 aggggagtg ttaggggggt taaggacggt gggggcgcca gctgggggt actgcacga 7620  
 ccttttacta aagccagttt cctggtctg atggtattg ctgagttatg ggagactaac 7680  
 cataggggag tggggatggg ggaacccgga ggtgtgcca tctttgccat gcccgagtg 7740  
 cctgggcagg ataagtctct agagatgcc acgtcctgat tcccccaac ctgtggagga 7800  
 aaccgcccc gcccagggc ctttgagggt gtgactctcg tgaggacct gaggtctggg 7860  
 65 atccttcggg actacctga gggccgaaa gtaatccagg ggttctggga agaggcggg 7920  
 agggagggtca gaggggggca gcctcaggac gatggaggca gtcagctcga ggctgaaaag 7980

5 ggagggaggg cctcgagccc aggcctgcaa gcgcctccag aagctggaaa aagcggggaa 8040  
 gggaccctcc acggagcctg cagcaggaag gcacggctgg cccttagccc accagggccc 8100  
 atcgtggacc tccggcctcc gtgccatagg agggcactcg cgctgccctt ctgacatgaa 8160  
 gtgtgtgggg atttgcagaa gcaacaggaa acccatgcac tgtgaatcta ggattatttc 8220  
 10 aaaaacaaagg ttacagaaa catccaagga cagggctgaa gtgcctccgg gcaagggcag 8280  
 ggagggcagc agtgatttta tttagctatt ttattttatt tacttacttt ctgagacaga 8340  
 gttatgctct tgtgccccag gctggagtgc agcggcatga tcttggctca ctgcaacctc 8400  
 ctgtctcttg gttcaagcaa ttctcgtgcc tcagcctccc aagtagctgg gatttcaggc 8460  
 gtgcaccacc acaccggcct aattttgtat ttttagtaga gatgggcttt caccatgttg 8520  
 gtcaagctga tctcaaaatc ctgacctcag gtgatccgcc caccctagcc tcccaaatgt 8580  
 ctgggattac aggcctgagc cactgcacct ggccatttta accattttta aacttccctg 8640  
 ggctcaagtc acaccactg gtaaggagt catggagttc aatttcccct ttactcagga 8700  
 gttaccctcc tttgataatt tctgtaattc ttcgtagact ggggatacac cgtctcttga 8760  
 15 catattcaca gtttctgtga ccacctgtta tcccatggga cccactgcag gggcagctgg 8820  
 gagggctgac gcttcagggt ccagtggggg tgcctctgc cagtagaaac ctgagtgtag 8880  
 atcagggcgc aagtgtggac actgtcctga atctcaatgt ctcagtgtgt gctgaaacat 8940  
 tagaaaatta aagtccatcc ctctactctt actgggatgg agccctctcc ctatccctcc 9000  
 ccaggggcag agggagttct ctactcctg tggagggaag aatgatactt tgttattttt 9060  
 cactgctggg actgaatcca ctgtttcatt tgttggtttg tttgttttgt tttgagaggc 9120  
 20 ggtttcactc ttgttgctca ggctggaggg agtgcaatgg cgcgactctg gcttacttga 9180  
 gcctctgcct ccaggttca agtgattctc ctgtctccgc ctcccatttg gctgggatta 9240  
 caggcacccc ccaccatgcc cagctaattt ttgtatttt tagtagagac ggggggtggg 9300  
 ggggttcacc atgttgcca ggctggcttc gaacttctga cctcagatga tccacctgcc 9360  
 tctgcctcct aaagtgttg gattacaggt gtgagccacc atgcccagct cagaatttac 9420  
 25 tctgtttaga aacatctggg tctgaggtag gaagctcacc ccactcaagt gttgtggtgt 9480  
 tttaaagcaa tgatagaatt tttttattgt tgttagaaca ctcttgatgt ttacactgt 9540  
 gatgactaag acatcatcag cttttcaaa agacactaac tgcaccata atactgggtg 9600  
 gctttctggg tatcagcaat cttcattgaa tgcggggagg cgtttcctcg ccactgacat 9660  
 ggtgttaatt actccagcat aatcttctgc ttccatttct tctcttccct cttttaaaat 9720  
 30 tgtgttttct atgttggtct ctctgcagag aaccagtgtg agctacaact taacttttgt 9780  
 tggaaacaa tttccaaacc gcccttttgc cctagtgcca gagacaattc acaaacacag 9840  
 ccccttaaaa aggcttaggg atcactaagg ggatttctag aagagcgacc tgtaactcta 9900  
 agtatttaca agacgaggct aacctccagc gagcgtgaca gccaggggag ggtgcgaggc 9960  
 35 ctgttcaaat gctagctcca taaataaagc aatttctctc gccagtttct gaaagtagga 10020  
 aaggttcaat ttaaggttgc gtttgttagc atttctagtgt ttgcccagct cagctacagc 10080  
 atccctgcaa ggcctcgga gaccagaag tttctcgccc ccttagatcc aaacttgagc 10140  
 aaccggagt ctggattcct gggaaagtcct cagctgtcct gcggttgtgc cggggcccca 10200  
 ggtctggagg ggaccagtgg ccgtgtggct tctactgctg ggctggaaat cgggctctct 10260  
 40 agctctgcag tccgaggctt ggagccaggt gectggacc cagggtgccc ctccaccttg 10320  
 tgcggggcgg atgtgaccag atgttgccct catctgccag acagagtgcc gggggcccag 10380  
 gtcaaggccg ttgtggctgg tgtgaggcgc ccggtgcgcg gccagcagga gcgcctggct 10440  
 ccatttccca ccctttctcg acgggaccgc ccggtgggt gattaaacaga tttggggtgg 10500  
 tttgctcatg gtggggacc ctgcgcgctt gagaacctgc aaagagaaat gacgggcttg 10560  
 45 tgtcaaggag ccaagtgcg ggggaagtgt tgcaggagg cactccggga ggtccgcgtg 10620  
 gcccgctccag ggagcaatgc gtctcgggt tgcctccag ccgcgtctac gcgcctccgt 10680  
 cctccccctc acgtccggca ttctgtgtgc ccggagcccc acgccccgcg tccggacctg 10740  
 gaggcagccc tgggtctccg gatcaggcca gcggccaaag ggtcgccgca cgcacctgtt 10800  
 cccagggcct ccacatcatg gccctctcct cgggttacc caccagctag gccgattcga 10860  
 50 cctctctccg ctggggccct cgctggcgtc cctgcacctt gggagcgcga gcggcgcgcg 10920  
 ggccggggaag cgcggcccag accccggggt ccgcccggag cagctgcgct gtccggggcca 10980  
 ggccggggtc ccagtggatt cgcgggcaca gacgcccagg accgcgtctc ccacgtggcg 11040  
 gagggactgg ggaccggggc acccgtctct ccccttacc ttccagctcc gcctctctcg 11100  
 cgcggaccct gccccgtccc gacccctccc gggtecccg cccagcccc tccggggcct 11160  
 cccagcccc ctccctctct tccgcggccc cgcctctctc tcgcggcgcg agtttcaggc 11220  
 55 agcgctgcgt cctgtgcgc acgtgggaag ccctggcccc ggccaccccc gcgatg 11276  
  
 <210> 4  
 <211> 104  
 <212> DNA  
 60 <213> Homo sapiens  
  
 <400> 4  
 gtgggcctcc ccggggtcgg cgtccggctg ggggttgagg cgccggggg gaaccagcga 60  
 catgcggaga gcagcgagg cgactcagg cgcttcccc gcag 104

<210> 5  
<211> 8616  
<212> DNA  
<213> Homo sapiens

5

&lt;400&gt; 5

10

15

20

25

30

35

40

45

50

55

60

65

gtgaggaggt ggtggccgtc gagggcccag gcccagagc tgaatgcagt aggggctcag 60  
aaaagggggc aggcagagcc ctggtcctcc tgtctccatc gtcacgtggg cacacgtggc 120  
ttttcgctca ggacgtcgag tggacacggg gatctctgcc tctgctctcc ctctgtcca 180  
gtttgcataa acttacgagg ttcacctca cgtttgatg gacacgagg ttccaggcgc 240  
cgaggccaga gcagtgaaca gaggaggtcg ggcgcggcag tggagccggg ttgccggcaa 300  
tggggagaag tgtctggaag cacagacgct ctggcgaggg tgccctgcagg ttacctataa 360  
tcctcttcgc aatttcaagg gtgggaatga gaggtgggga cgagaacccc ctcttctctg 420  
gggtgggagg taagggtttt gcaggtgcac gtggtcagcc aatatgcagg tttgtgttta 480  
agatttaatt gtgtgttgac ggcaggtgac ggtggtcac gccggtaatc ccagcacttt 540  
gggaagctga ggcaggtgga tcacctgagg tcaggagtgt gagaccagcc tgaccaacat 600  
ggtgaaaccc tatctgtact aaaaatacaa aaattagctg ggcagtgttg tgtgtgctg 660  
taatccagc tacttgggag gctgaggcag gagaatcact tgaaccagg aggcggaggc 720  
tgcagtgaac tgagattgtg ccattgtact ccagcctggg cgacaagagt gaaactctgt 780  
ctttaaaaaa aaaaagtgtt cttgtattgt gccaggacag ggtagagggg gggagataag 840  
actgttctcc agcacagatc ctggtcccat cttaggtat gaagagggcc acatgggagc 900  
agaggacagc agatgggtcc acctgctgag gaaggagacag tgtttgtggg tgttcagggg 960  
atggtgctgc tgggcccctg cgtgtcccca ccctgtttt ctggatttga tgttgaggaa 1020  
cctccgctcc agccccctt tggctcccag tgcctccagg cctaccgtg gcagctagaa 1080  
gaagtcccg tttcaccccc tccccacaaa ctcccaagac atgtaagact tccggccatg 1140  
cagacaagga gggtgacctt cttggggctc tttttttct ttttttctt ttatggtggc 1200  
aaaagtcata taacatgaga ttggcactcc taacaccgtt tctgtgttac agtcagaaat 1260  
tgctaactcg gcggtgttta cagcaggttg cttgaaatgc tgcgtcttgc gtgactggaa 1320  
gtccctaccc atcgaaacgc agctgcctca cactgtctgc ggtcaggtg gaccacgccg 1380  
agtcagataa gcgtcatgca acccagtttt gctttttgtg ctccagcttc cttcgttgag 1440  
gagagtttga gttctctgat caggactctg cctgtcattg ctgttctctg acttcagatg 1500  
aggtcacaa ctgcccctgg cttatgcagg gagtgcaggc tggctcccgg gtgtccctgt 1560  
cacgtgcagg gtgagtgagg cgttgcccc aggtgtccct gtcacgtgta gggtagtgta 1620  
ggcgcgcccc ccgggtgtcc ctgtcccgtg cagcgtgatt gaggtgtggc ccccggtgtg 1680  
ccctgtcacg tgtagggtag gtgaggcgcc atccccgggt gtccctgtca cgtgtagggt 1740  
gagtgaggcg tgggtcccgg gtgtccctgt cccgtgcagg gtgagtgagg cactgtcccc 1800  
gggtgtccct gtcacgtgca gggtagtgga ggcgcggctc ccgggtgtcc ctctcaggtg 1860  
tagggtgagt gaggcgcggc cccaggggtg cctgttcacg tgtagggtag gtgaggcacc 1920  
gtcccctggg gtcccctcca ggtatagggt gagtgcaggc ctgtcccgg gtgtccctgt 1980  
cacgtgcagg gtgagtgagg cgcggccccc ggggtgtccct ctccaggtgca gggtagtgta 2040  
ggcgctgtcc ctgggtgtcc ctgtctctgt tagggtagt gaggtctctg cccaggtgtg 2100  
ccttggcggt tgtcacttg agcttgcctc tgaatgtttg ctcttctat agccacagct 2160  
gcgcgggttg cccattgcct gggtagatgg tgcaggcgca gtgctgttcc ccaagcctat 2220  
cttttctgat gctcggtctc tcttgggtcac ctctccgttc cattttgcta cggggacacg 2280  
ggactgcagg ctctcgctcc cgcgtgtcca ggcactgcag ccacagcttc aggtccgctt 2340  
gcctctgttg ggcctggctt gctcaccacg tgcccgccac atgcatgctg ccaatactcc 2400  
tctcccagct tgtctcatgc cgaggctgga ctctgggtg cctgtgtctg ctgccacgtg 2460  
ttgctggaga catcccagaa agggttctct gtgcccgtga ggaaagcaag tcaccccagc 2520  
ccccctcact gtccctgttt ctcccaagct gccctctgct tggccccct tgggtgggtg 2580  
gcaacgcttg tcaacttatt ctgggcacct gccgtcatt gcttaggctg ggtctgtcct 2640  
ccagtcgccc cctcacatgg attgaagtc agccacaggt tggagtgtct ctgtctgtct 2700  
cctgctctga gacccacgtg gagggccggg gtctccgcca gccttctgca gacttccctc 2760  
ttgggtctta gttttgaatt tcactgattt acctctgacg tttctatctc tccattgtat 2820  
gctttttctt ggtttattct ttcattcctt tcttagcttc ttagtttagt catgccttct 2880  
cctctaagt ctgccttacc tgcacccctg gttttgatgt gaagtaatct caacatcagc 2940  
cactttcaag tgttcttaaa atacttcaaa gtgttaatac ttcttttaag tattcttatt 3000  
ctgtgatttt tttctttgtg cacgtgtgt tttgacgtga aatcattttg atacagtgta 3060  
cttttaagta ttctttagct tattctgtga ttcttttgag cagtgcgtta tttgaacact 3120  
gtttatgttc aagatatgta gagtatcaag atacgtagag tatttttaag tatcatttta 3180  
ttattgattt ctaactcagt tgtgtagtg tctgtataat accaattatt tgaagtgtgc 3240  
ggagccttgc tttgtgatct agtgtgtgca tgggtttccag aactgtccat tgtaaatttg 3300  
acatccgtgc aatagtgggc atgcatgttc actatatcca cttatttaag gtccagtgca 3360  
aagcttctgt ctccctctag atgcataaaa ttccaagaag gaggccatag tccctcacct 3420  
gggggatggg tctgttcatt tcttctcgtt tggtagcatt tatgtgaggc attgttaggt 3480  
gcatgcacgt ggtagaattt ttatcttctc gatgagtga tcttttgag acttctatgt 3540

ctctagtaaat ctagtaattc tttttttaa ttgctcttag tactgccaca ctgggcttct 3600  
 ttgtgattagt attttctctgc tgtgtctgtt ttctgccttt aatttatata tatatatata 3660  
 tttttttttt ttttgagaca gagtcttggc ctgtcgccca gggtagtgac agtgggtgtga 3720  
 5 tcacagggtca gtgtaacttt taccttctgg cctgagccgt cctctcacct cagcctcctg 3780  
 agtagctgga actgcagaca cgcaccgcta cacttgcta atttttaaat tttttctgga 3840  
 gacagggtct tgcgtgtgtg cccaggctgg tctcaaactc ttggactcaa gggatccatc 3900  
 taccttcggct tcccaaagt ctgaattaca ggcacgagcc accatgtctg gcctaatatt 3960  
 caacactttt atattcttat agtgtgggta tgcctgttta acagcatgta ggtgaatttc 4020  
 caatccagtc tgacagtcgt tgtttaactg gataacctga tttattttca tttttttgtc 4080  
 10 actagagacc cgcctgggtgc actctgattc tccacttgcc tgttgcatgt cctcgttccc 4140  
 ttgtttctca ccacctcttg ggttgccatg tgcgtttcct gccgagtggt tgttgatcct 4200  
 ctctgtgctt cctgggtcact gggcatttgc ttttatttct ctttgcctag tgttaccctc 4260  
 tgatcttttt attgtcgttg tttgcttttg tttattgaga cagtctcact ctgtcaccca 4320  
 ggtctggagt taatggcaca atctcggtc actgcaacct ctgcctcctc ggttcaagca 4380  
 15 gttctcattc ctcaacctca tgagtgcgt ggaattacagg cgcaccacac cagcctcctg 4440  
 taatttttgt atttttagta gagataggct ttcaccatgt tggccaggct ggtctcaaac 4500  
 tcttgacctc aagtgatctg cccgccttgg cctccacacg tgctgggatt acagggtgcaa 4560  
 gccaccgtgc ccgcataacc ttgatctttt aaaatgaagt ctgaaacatt gctacccttg 4620  
 tcttgagcaa taagaccctt agtgratttt agctctggcc acccccagc ctgtgtgtgt 4680  
 20 ttttccctgc tgacctagt ctactctcagg catcttgaca ccccacaaag ctaagcatta 4740  
 ttaattattg tttccgtgtt gagtgtttct gtagctttgc ccccgccctg cttttctctc 4800  
 tttgttcccc gtctgtcttc tgcctcaggc ccgcgtctg ggggtccctt ccttgcctt 4860  
 tgcgtgggtc tctgtcttg ttattgtctg taaacccag ctttacctgt gctggcctcc 4920  
 atggcatcta gcgacgtccg gggacctctg cttatgatgc acagatgaag atgtggagac 4980  
 25 tcacgaggag ggcggtctt tggccctgt agtgtctgga gcaccacgtg gccagctgtc 5040  
 ctttagccagt gagtgcagc aacgtccgct cggcctgggt tcagcctgga aaacccagg 5100  
 catgtcgggg tctgtgtgct ccgcgtgtc gagtgtgaaa tcgcgcaaac ctgcggtgtg 5160  
 gcgcagctc tgacgggtgt gcttggcggt gtagtgcctg cttctctcct tctgtctggg 5220  
 aaccaggaca aaggatgagg ctccgagccg ttgtcgccca acaggagcat gacgtgagcc 5280  
 30 atgtggataa ttttaaaatt tctaggctgg gcgcggtggc tcacgcctgt aatcccagca 5340  
 ctttgggagg ccaaggcggt tggatcacga ggtcaggagg tcgagaccat cctggccaac 5400  
 atgatgaaac cccatctgta ctaaaaaac aaaaattagc tggcggtgtg ggcgggtgcc 5460  
 tgtaatccca gctactcggg aggtctgagg aggagaattg cttgaacctg gtagtggaa 5520  
 gttgcagtg gcccacattg caccactgca ctccagcctg gcaacacagc gagactctgt 5580  
 35 ctcaaaaaaa aaaaaaaaaa aaaaaaaaaa aattctagta gccacattaa aaaagtaaaa 5640  
 aagaaaaggt gaaattaatg taataataga ttttactgaa gccacagcat tccacacctc 5700  
 atcatttttag ggtgttattg gtgggagcat cactcacagg acatttgaca ttttttgagc 5760  
 tttgtctcgc ggtatccctg tgtaggctcc gtgcgtggcc atctcggcct ggacctgtct 5820  
 ggcttcccat ggccatggct gttgtaccag atgggtcagg tccgggatga ggtcgccagg 5880  
 40 cctcagtgga gctggatctg cagtgtccgg atgggtcacg tctgggatga ggtcgccagg 5940  
 ccctgctgtg agctggatgt gtggtgtctg gatggtgag gtccagggtg aggtctccag 6000  
 gccctcgggt agctggaggt atggagctcg gatgatgag gtccgggggt aggtcgccag 6060  
 gccctgctgt gagctggatg tgtgtgtctt ggaatgtgca ggtcagggtg gaggtctcca 6120  
 45 ggccctcggg aagctggagg tatggagtcc ggaatgagca ggtccgggtg gaggtcgcca 6180  
 ggccctgctg tgagctggat gtgtgtgtct tggatggtg aggtctgggg tgaggtcacc 6240  
 aggcctcgcg gtgagctggg tgtgcggtgt ctggatgggt caggctctgga gtgaggtcgc 6300  
 cagacggtgc cagaccatgc ggtgagctgg atatgcggtg tccggatggt gcaggtctgc 6360  
 ggtgaggttg ccaggccctg ctgtgagttg gatgtgggt gtccggatgc tgcagggtcc 6420  
 gtgtgaggtc accaggccct gctgtgagct ggaatgtgtg tgtctggatg gtgaggtct 6480  
 50 ggggtgaagg tcgccaggcc cctgcttctg agctggatgt gtgggtgtct gatggtgcag 6540  
 gtctggagtg aggtcgccag gccctcgggt agctggatgt gcaggtgtcc gatggtgcag 6600  
 gtccgggggt aggtcgccag accctgcggt gagctggatg tgcggtgtct ggaatgtgca 6660  
 ggtctggagt gaggtcgcca ggccctcgggt gagctggatg tatggagtcc ggaatgtgca 6720  
 ggtccgggggt gaggtcgcca gacctgctg tgagctggat gtccggtgtc tggatggtac 6780  
 55 aggtctggag tgaggtcgcc agacctgct gtgagctgga tatgcggtgt ccggatggtg 6840  
 caggtcagggt gtgaggtctc caggccctcg gtgagctgga ggtatggagt ccggatggtg 6900  
 caggtccggg gtgaggtcgc caggccctgc tgtgaactgg atgtcgggcg tctggatggt 6960  
 gcaggtctgg ggtgtggtcg ccaggccctc ggtgagctgg aggtatggag tccggatgat 7020  
 gcaggtcccg ggtgaggtcg ccaggccctc ctgtgagctg gatgtgagg gtctggatg 7080  
 60 tgcaggtctg ggggtgtggtc gccaggccct cgtgtgagct gaggtatgga gtccggatga 7140  
 tgcaggtccg ggggtgaggt gccaggccct gctgtgagct ggaatgtgct tatccggatg 7200  
 gtgcaggtccg ggggtgaggt gccaggccct gctgtgagct ggaatgtgct tatccggatg 7260  
 gtgcaggtct ggggtgaggt caccaggccc tgcggtgagc tgggtgtgct gtgtccgggt 7320  
 gctgcaggtc cgggtgaggt tcgccaggcc ctgcgtgagc tggatgtgct gtgtcccggt 7380  
 65 gtccggatgg tgcaggtcca ggggtgaggt gctaggccct tgggtgggct gatgtgctg 7440

gtccggatgg tgcaggctctg ggggtgaggtc gccaggccctt tggtagagctg gatgtgcgggt 7500  
 gtctgcagtg tgcaggctctg ggggtgaggtc gccaggccctt tggtagagctg gatgtgcgggt 7560  
 gtccggatgg tgcaggctccg gcgtgaggtc gccaggccctt gctgtgagct ggatgtgcgg 7620  
 5 tgcctggatg gtgcaggctcc ggggtgaggt agccaaggcc ttcggtagagc tggatgtggg 7680  
 gtgtccggat ggtgcaggctc cggggtgagg tcgccaggcc ctgaggcttag ctggatatgc 7740  
 ggtgtccgga tgggtgcaggc cggggtgag gtccaggcc cctgaggcta gctggatgtg 7800  
 cgggtgtctgg atggtgcagg tccgggtgga ggtcgccagg cctgctgtg agctggatgt 7860  
 gctgtatccg gatggtgcag gtccgggtg aggtcgccag gccctgcagt gagctggatg 7920  
 10 tgcgttatcc ggaagggtgca ggtctggcgt gaggtcgcca ggccctgagg ttagctggat 7980  
 atgcgggtgc ggaagggtgca ggtccgggtg gaggtcacca ggccctgagg ttagctggat 8040  
 gtgcgggtgc cggatggtgc aggtctgggg tgaggtcgcc aggcctgctg gtgagctgga 8100  
 tgcgtctgat ccggatggtg cagggtccggg gtgaggtcgc caggccctgc ggtgagctga 8160  
 atgtgctgta tccggatggt gcaggctctg cgtgaggtcg ccaggccctg cggtagagctg 8220  
 gatgtgcagt gtacggatgg tgcaggctcc ggtgaggtc gccaggccct cggtagggct 8280  
 15 gtatgtgtgt tgcctggatg gtgcaggctc ggggtgagtt cggcaggccc tgcggtagagc 8340  
 tggatgtgtg gtgtctggat gctgcagggt cggggtgagt tcgccaggcc ccggtgagc 8400  
 tggatatgcg gtgtcccggt gtccgaatgg tgcaggctca gggtagaggt gccaggccct 8460  
 tggtagggctg gatgtgcctg gtccggatgg tgcaggctct gggtagaggt gccaggccct 8520  
 20 tggtagagctg gatgtgcggg gtccggatgg tgcaggctcc gggtagaggt accaggccct 8580  
 cggtagatctg gatgtggcat gtccctctctg ttttaag 8616

<210> 6  
 <211> 2089  
 <212> DNA  
 25 <213> Homo sapiens

<400> 6  
 gtactgtatc cccacgccag gcctctgctt ctggaagtc tggaaacacca gcccgccctc 60  
 agcatgcgcc tgtctccact tgcctgtgct tccctggctg tgcagctctg ggcctgggagc 120  
 30 caggggccccc gtccacaggcc tgggtccaagt ggattctgtg caaggctctg actgacctgga 180  
 gctcacgttc tcttacttgt aaaatcagga gtttgtgcca agtggctctt aggggttgtg 240  
 aagcagaagg gatttaaat agatggaaac actaccacta gcctccttgc ctttccctgg 300  
 gatgtcgggtc tgattctctc tctctttttt ttttctttt tgagatggag tctcactctg 360  
 35 ttgcccgcc tggagtgccag tggcataatc ttggctcact gcaacctcca cctcctgggt 420  
 ttaagcgatt caccagcctc agcctcctaa gtacgtggga ttacaggcac ctgccaccac 480  
 gcctgggctaa tttttgtact tttaggagag acgggggttt accatgttgg ccaggctggt 540  
 ctcgaaactca tgacctcagg tgatccaccc accttgccct cccaaagtgc tgggtttaca 600  
 ggctaagcca ccgtgccag ccccgattc tcttttaatt catgctgttc tgtatgaatc 660  
 40 ttcaatctat tggatttagg tcatgagagg ataaaatccc acccacttgg cgactcactg 720  
 cagggagcac ctgtgcaggg agcacctggg gataggagag ttccaccatg agctaacttc 780  
 taggtggctg catttgaatg gctgtgagat ttgtctgca atgttcggct gatgagagtg 840  
 tgagattgtg acagattcaa gctggatttg catcagtgag ggacgggagc gctggtctgg 900  
 gagatgccag cctggctgag cccaggccat ggtattagct tctcctgtgc ccgccaggc 960  
 45 tgactgtgga gggctttagt cagaagatca gggcttcccc agctccccct cactctcag 1020  
 tccctggggg gccttgtgac accccatgcc ccaaatcagg atgtctgcag agggagctgg 1080  
 cagcagacct cgtcagaggt aacacagcct ctgggctggg gaccccgagc tgggtctggg 1140  
 gccatttcct tgcattctgg gagggtcag ggccttccct gtgggaacaa gttataacac 1200  
 aatgcacctt acttagactt tacacgtatt taatgggtgt cgacccaaca tgggtcattg 1260  
 50 accagtatct tggaaagaat ttaattgggg tgaccggaag gagcagacag acgtggtgg 1320  
 cccaagatg ctcttctgca ctactgggac tgtgtgtctg cctggggggc ctggaggcc 1380  
 cctcctcccc ggacagggtc ccgtgccttt tctactctgc tgggctgagc gctgagggc 1440  
 agggcaccag ctccggagca ccccgggccc cagtgtccac ggagtgccag gctgtcagcc 1500  
 acagatgccc aggtccagggt gtggccgctc cagccccctg gccccatgg gtgggttttg 1560  
 55 gggaaaaggc caagggcaga ggtgtcagga gactggtggg ctcatgagag ctgattctgc 1620  
 tccctggctg agctgcccct agcagcctct ccgcccctct ccatctgaag ggatgtggct 1680  
 ctttctacct gggggctcct cctggggcca gccttgggt accccagtgg ctgtaccaga 1740  
 gggacaggca tcctgtgtgg aggggcatgg gttcacgtgg cccagatgc agcctgggac 1800  
 caggctcccc ggtgtctgat gtgggacagt caccctgggg gttgaccgcc ggactgggac 1860  
 60 tccccagggt tgaactatag accaggtgtc cagggtgccc gcaagtagag gggctctcag 1920  
 aggcgtctcg ctggcatggg tggacgtggc cccgggcatg gccttcagcg tgtgtcggc 1980  
 tgggtgcctt gagccctcac tgagtgggtg ggggcttctg gcttccctg agcttcccc 2040  
 tagtctgttg tctggctgag caagcctcct gaggggctct ctattgcag 2089

10 / 18

<210> 7  
<211> 687  
<212> DNA  
<213> Homo sapiens

5

<400> 7  
gtggctgtgc tttggtttta cttccttttt aaacagaagt gcgttttgagc cccacatttg 60  
gtatcagctt agatgaaggg cccggaggag gggccacggg acacagccag ggccatggca 120  
cggcgccaac ccatttgtgc gcacagttag gtggccgagg tgccgggtgcc tccagaaaag 180  
10 cagcgtgggg gtgtaggggg agctcctggg gcaggagacag gctctgagga ccacaagaag 240  
cagccggggc agggcctgga tgcagcacgg cccgaggtcc tggatccgtg tcctgctgtg 300  
gtgcgcagcc tccgtgcgt tccgtttacg gggcccgggg accaggccac gactgccagg 360  
agcccaccgg gctctgagga tcctggacct tgccccacgg ctctctgcacc ccaccctgt 420  
ggctgcggtg gctgcgtga ccccgctatc tgaggagagt gtgggttgag gtggacagag 480  
15 gtgtggcatg aggatcccg gtgcaacaca catgcggcca ggaaccctt tcaaacaggg 540  
tctgaggaag ctgggagggg ttctagggtc cgggtctggg tggctgggga cactggggg 600  
gggtctgttc tcccctgggt ccctatggtg ggggtgggac ttggccggat ccactttcct 660  
gactgtctcc catgctgtcc ccgccag 687

20

<210> 8  
<211> 494  
<212> DNA  
<213> Homo sapiens

25

<400> 8  
gtgggtgccc gggacccccg tgagcagccc tgctggacct tgggagtggc tgccctgattg 60  
gcacctcatg ttgggtggag gaggtactcc tgggtgggccc gcaggagtg caggtgaccc 120  
tgtcactgtt gagacacac ctggcaccta ggttgaggc cttcagcctt tcctgcagca 180  
catggggccc actgtgcacc ctgactgccc gggctcctat tcccaaggag ggtcccactg 240  
30 gattccagt tccgtcagag aaggaaaccg aacggctcag ccaccaggcc ccggtgcctt 300  
gcaccccaag cctgagccag gggctcctg tcctgaggct cagagagggg acacagcccc 360  
ccctgccctt ggggtcttga gtgtggggg tcagagagag agtgggggac accgccaggc 420  
caggccctga gggcagaggt gatgtctgag tttctgcgtg gccactgtca gtctcctcgc 480  
ctccactcac acag 494

35

<210> 9  
<211> 865  
<212> DNA  
<213> Homo sapiens

40

<400> 9  
gtaagggtca cgtgtgatag tctgtctccag gatgtgtgtc tctgggatat gaatgtgtct 60  
agaatgcagt cgtgtctgtg atgcgtttct gtgggtggagg tacttccatg atttacacat 120  
ctgtgatatg cgtgtgtggc acgtgtgtgt cgtgggtgcat gtatctgttg cgtgcatatt 180  
45 tgtgtgtgtg gtgtgtgtgg cacgtgtgtg tccatggtgt gtgtgcctgt ggtgtgcatg 240  
tgtgtgtgtc tgtgacacgt gcatgttcat gctgtgtgct gcatgtctgt gatgtgctta 300  
tttgtgtgtg gtgtgtgcat gtgtccgtga catatgcgtg tctatggcat ggggtgtgtg 360  
ggccccctgg ccttactcct tccctcctca ggcatggtcc gcaccattgt cctcacgctc 420  
tcgggtgctg gtttggggag ctccacattc agggctcctca cttctagcat ggggtgccct 480  
50 gtcctgtcac agggctgggg cttgggagact gtaagccagg tttgagagga gagtagggat 540  
gctgtgtgta ccttccctgga cccctggcac cccagggacc ccagtctggc ctatgccggc 600  
tccatgagat ataggaaggc tgattcaggc ctgcgtcccc gggacacact cctcccagag 660  
cggccggggg ccttggggct cggcaggggt gaaaggggccc ctgggcttgg gttcccaccc 720  
agtgtgcatg agcacgctgg aggggttaag cctcaaagtc gtgccaggcc ggggtgcaga 780  
55 ggtgaagaag tatccctgga gcttcgggtc ggggagaggc acatgtggaa acccaagaag 840  
acctctttct ctgactttct gagct 865

60

<210> 10  
<211> 3782  
<212> DNA  
<213> Homo sapiens

65

<400> 10  
tgtgggattg gttttcatgt gtgggatagg tgggatctg tgggattggt ttttatgagt 60  
ggggaacac agagttcaag gcgagcttcc ttcctgtagt gggctctgag gtgctccaac 120

agcttttattg agggagaccat atcttctctt gaactatggt cgggtttata gtaagtcagg 180  
 ggtgtggagg cctccctcgt gctccctgtt ctgttctctt cactctgggg tcgtgtgggt 240  
 cctgtctgtg tgtgtggcgg gtgggcagggt cttccaggcc tcttctgtgt cattggcctg 300  
 gatgtggccc tggctacgct ccgtccttgg aattccctcg cgagtggagg gctttctctt 360  
 5 tttctttttt tctttctttt tttttttttt tgataacaga gtctcgtctt ttttgccca 420  
 ggctggagtg gtttggcgtg atcttggctc actgcaacct gtgcttcttg agttcaagca 480  
 attctcttgc ctacagcctc caagtagctg gaattatagg cgcccaccac catgctgact 540  
 aatttttcta atttttagtag agacgaggtt tctccatgtt ggccagggtg gtctcgaact 600  
 10 cctgacctca ggtgatcctc ccacctcggc ctcccaaagt gctgggatga caggtgtgaa 660  
 ccgcccggcc cggccgagac tcgttctctg cagcttccgt gagatctgca gcgatagctg 720  
 cctgcagcct tgggtctgac aacctccgtt tctcttctcc aggtctcgtt aggggtcttt 780  
 ccatttcatg actctcttca cagaagagtt tcacgtgtgc tgatttcccg gctgtttctt 840  
 gcgtaattgg tgtctgtgtt ttatcgatgg cctccttcca tttctttag gcttgttta 900  
 ttgttgtttt tccggctcct tgaaggaaaa gtctcgatta tggatgtttg aacttctttt 960  
 15 tctaaacaag catctgaagt tgccgttttc cctctaaagc agggatcccc agggcccttg 1020  
 ctgtggagtg gcaccggtct ggggctgtt aggaacccgg cgacacagg gaggtatagg 1080  
 ggggtgtggg gagccagcgt tcccgcttga gcccgccttc tctcagatca gcagtggcat 1140  
 gcggtgtctca gaggcgcaca caccctactg agaactgtgc gtgagagggg tctagattct 1200  
 20 gtgtctctta tgggaatcta atgectgatg atctgaggtg gaaccgtttg ctcccaaac 1260  
 catcccttcc cccactgctg tctgtggaa aaatcgtctt ccacgaaacc agtccctggg 1320  
 accacaatgg ttggggaccc tgtgctaaag acctgcttca gcagcctctc gtcagtgttg 1380  
 atatatggc ttttctgtgt tgagtcacga ataattacgg atttctgtga tgcttctcgc 1440  
 cgacctcaga cccatgggct atttgtgggc gtgttgcctg ctctgggtt gggaaaggtg 1500  
 25 caggcccat gtaccttctt gttactgctt tccagggttg tctcagggt tgaaatcgta 1560  
 tcgatgtggt tttagccctc ggccctgccc ccagctcctg ggggctggg aacatgtcga 1620  
 agcacagagt caccgtgccc gtcttttgat gccctacaag ctccaggcct cctgtgtccg 1680  
 tgttagtggt tgtcacgtgc ctgctcacat cctgtcttgg ggacgcagg gcttagcagg 1740  
 tcccgtagta aatgacaagc gtcttggggg agtctgcaga ataggaggtg ggggtgcccg 1800  
 30 tctctctccc gcgtcttcag actcttctcc tgccgtgtgt gtggctgcac ctgcatcccc 1860  
 gcaatccctc cagcactggg ctggagaggg ccgggagctc gagtgccact tgtgccactg 1920  
 gacttgggat ggacgtcggc caccgggggtc tgatgtgtgg tgactgtgga tggcggttg 1980  
 tcacaggggt ctgatgtgtg gtgactgtgg atggcggtcg tggggtctga tgtgtgact 2040  
 gtggatggcg gtctgtgggt ctgatgtgtg gtgactgtgg atggcggtcg tggggtctga 2100  
 35 tgtgttgact gtggatggcg gtctgtgggt ctgatgtgtg gactgtggat ggcggtcggt 2160  
 ggggtctgat tgggtactgt ggatggcagt cgtgggggtc gatgtgtggg gactgtggat 2220  
 ggcggtcgtg ggggtctgat tgggtactgt ggatggcagt cgtgggggtc gatgtgtggg 2280  
 gactgtggat ggcggtcgtg ggggtctgat tgtgtgtact gtggatggcg gctgtgggtg 2340  
 40 ctgatgtgtg gtgactgtgg atggcggtcg tggggtctga tgtgtgtgga ctgtgtggat 2400  
 cgggtcgtgg gtctgatgtg gtgactgtgg atggcggtcg tggggtctga tgtgtgtgga 2460  
 ctgtgtgatg tgatcgttca caggggtctg atgtgtgtg actgtggat ggcggtcgtg 2520  
 ggtctgatgt gtggtgactg tggatgggtga tcgggtcacag ggggtctgat tgtgtgact 2580  
 gtggatggcg gtctgtgggt ctgatgtgtg gtgactgtgg atggcggtcg gtcggggggg 2640  
 tctgatgtgt ggtgactgtg gatggcgatc ggtcacagg gtctgatgtg tgggtactgt 2700  
 45 ggatggcggt cgtgggggtc gatgtgtgt gactgtggat ggcggtcgtg ggggtctgat 2760  
 tgtgtgtact gtggatggcg gtctgtgggt ctgatgtgt gactgtggat ggcggtcgtg 2820  
 ggggtctgat tgggtactgt ggatggcggt cgtgggggtc gatgtgtgt gactgtggat 2880  
 ggcggttggg cccgggggtc tgatgtgtg tgactgtgga tggcggtcgt ggggtctgat 2940  
 gtgggtgact tggatggcag tcgtgggggt tgatgtgtg tgactgtgga tggcggtcgt 3000  
 50 ggggtctgat gtgtgggtgac tgtggatggc ggtcgtgggt tctgatgtgt ggtgactgtg 3060  
 gatggcggtc gtgggggtct atgtgtgtgt actgtggatg gcgggtcgtg ggtctgatgt 3120  
 ggtgactgtg gatggcggtc gtgggggtct atgtgtgtgt actgtggatg gtgatcgggtc 3180  
 acaggggtct gatgtgtgtg gactgtggat ggcggtcgtg ggggtctgat tgtgtgact 3240  
 gtggatggcg gtctgtgggt ctgatgtgtg gactgtggat ggcggtcgtg ggggtctgat 3300  
 55 tgtgtgtact gtggatggcg gtctgtgggt ctgatgtgtg gtgactgtgg atggcagtcg 3360  
 gtcacagggt tctgatgtgt ggtgactgtg gatggcggtc gtgggggtct atgtgtgtg 3420  
 actgtgggat gcgggtcgtg ggtctgatgt gtggtgactg tggatggcggt cgtgggggtc 3480  
 tgatgtgtgt tgactgtgga tggcggtcgt gggtctgat gtggtgactg tggatgggtg 3540  
 tcgggtcacag ggggtcgtgat tgtgtgtagt gcaggtggag tcccgaggtg gtctgtagt 3600  
 60 actttgcgtc ctccggcccc cgccccctg tcccaaaaca gaagcttccc aggcgctctc 3660  
 tgggtctcat cccgccatcg ggcttggcgg caggtccaca cgtcctgatc ggaagaaaca 3720  
 agtgccacag tctggccggg gcaggccaca ttgtgtgtc atgccctctc ctctgccggc 3780  
 ag 3782

12 / 18

<210> 11  
<211> 980  
<212> DNA  
<213> Homo sapiens

5

<400> 11  
gtctgggcac tgccttgac gggtgggcac ggactcccag cagtgggtcc tccccgggc 60  
aatcactggg ctcctgaccg gacagactgt tggccctggg gggcagtggg gggaaatgagc 120  
tgtgatgggg gcatgatgag ctgtgtgccc tggcgaaatc tgagctgggc catgccaggc 180  
10 tgcgacagct gctgcattca ggcacctgct cacgtttgac tgcgcgccct ctctccagtt 240  
ccgcagtgcc ttgtttcatg atttgctaaa tgtcttctct gccagttttg atcttgaggc 300  
caaaggaaaag gtgtccccct ccttttaggag ggcaggccat gtttgagccg tgtcttgccc 360  
agctggcccc tcagtgtctg gtctgaggcc aaaggaaacg tgtccccctt cttaggagga 420  
cgggcccgtgt ttgagccacg ccccgctgag cgggcctctc agtgctgggt ctgtccacgt 480  
15 ggccctgttg cctcttgacg atgtggtctg tccacgtggc cctgtggctc ttgacagatg 540  
cctgttagca cttgctcggc tctaggggac agtcgtgtcc accgcatgag gctcagagac 600  
ctctgggcga atttctcttg ctcccagggt ggggggtggag gtggcctggg ctgctgggac 660  
ccagaccctg tgcccggcag ctgggcagca actcctggat cacatatgcc atccgggcca 720  
cgggtggctg tgtgggtgtg agcccagctg gaccacaggg tggcccagag gagacgttct 780  
20 ggtcacaca ctctgcctaa gcccattgtg gtctgcagag actcggcccg gccagccacc 840  
gatggccctg cattccagcc cagccccgca cttcatcaca aacactgacc ccaaaaggga 900  
cggagggtct tggccacgtg gtctgacctg tctcagcacc caccggctca ctcccatgtg 960  
tctcccgctc gctttcgacg 980

25

<210> 12  
<211> 2485  
<212> DNA  
<213> Homo sapiens

30

<400> 12  
gtgagtcagg tggccagggt ccattgccct gcgggtggct gggcgggctg gcagggcttc 60  
tgctcacctc tctctgccc ctccccact gnccttctgc ccggggccac cagagctccc 120  
ttttctggcc cccgccccct ccggctcctg ggctgcaggc tcccagggcc ccggaaacat 180  
35 ggctcggctt gcggcagccg gagcgagca ggtgccacac gaggcctgga aatggcaagc 240  
ggggtgtgga gttgctcctg cgtggaggac gaggggcggg ggggtgtgtc gggtcagggt 300  
tgcgcgagc gtttgagcct gcagctgtgc agctccaaat tactactgac gctggacacc 360  
cggctctcac acgcttgat ctctctctcc cgatacaaaa ggattttatc cgattctcat 420  
tctgtccct gtctgtgtac ccccgcgagg gcgcgggtct tctctctgt gactagattt 480  
cccatctgga aagtgcgggg ttgaccgtgt agtttgctcc tctcgggggg cctgtgttg 540  
40 ccatggggca ggcggccttg gagagctgac gtacacagc cactgggtga gccacactca 600  
cgttggtaga gccacagtgc ctggtgccac atcacgtcct ctggatttta agtaaaacca 660  
cacacctccc ggcaggcatc tgcctgagc cctgtgtgtg cctggggaga gtggtagcac 720  
ggaggaatt cgtgcacact caaggtcatc agcaaggcca tccgcagcca ggtggaacct 780  
ggaggcctct ctctgggac gtctccagcg gataaaggac tgtgcacagc ttcggaagct 840  
45 ttattttaa aatataacta ttaattattg cattataagt aatcactaat ggatcagca 900  
attataatat ttattaaagt ataattagaa atattaagta gtacacacgt tctggaaaaa 960  
cacaaattgc acatggcagc agagtgaatt ttggccgagg gacacgtgtg cacatgtgtg 1020  
taagcggccc ccaggccccc agaattcgct gacaaaagca cctccccaga gaagccacca 1080  
cgggctcctc tctgtgtctg gaattttatt aagatggatc aagtcacgta ccgtccacgt 1140  
50 gtggcagggc tttggggaat gtgaggtgat gactgcgtcc tcatgccctg acagacagga 1200  
gggtgactgt tctgtcctgt ccctaggaca cggacaggcc cgaagctcta gtccccatcg 1260  
tgggtccagtt tggcctctga ataaaaacgt cttcaaaacc tgttgcccca aaaactaaga 1320  
acagagagag tttcccatcc catgtgctca caggggcgta tctgcttgcg ttgactcgct 1380  
gggctggccg gactcctaga gttgtgtcgt gtgcttctgt gcaaaaagtg cagtcctctt 1440  
55 gcccatacct gtgatatctg caccagcaag gaaagcctct tttcttttct tctttttttt 1500  
ttttttgaga cggaaacgtca ctgtgtctg cctgggcttg agtgacagtg cgcgatctca 1560  
actcactgca acctccgcct cccgggttcc agcatttctc ctgcctcagc ccccagca 1620  
gctgagatta caggcaccca cccctgccc ctggctaatt tttgtatttt tagtagagag 1680  
gggtttttgc catgttgccc aggtgtgtct cgaactcctg acctcaggtg atccaccac 1740  
60 ctccgctccc caaagtgtg ggattacagg tgtgagccat cacgcccagc cggaaagcct 1800  
ctttttaagg tgaccacctc tagcgttccc gcaaaaatac aggtcttgtt tttgcagtag 1860  
gctgcaagcg tctcttagca acaggagtgg cgtcctgtgg gctctgggga tggctgaggg 1920  
tcgctgtgca gccatgctct ctgtgtgcac ctttaggttc cacgggggcta tctgtctctc 1980  
actgtttgtc tgaaaacgca cccttggcat ccttgttttg agagtttctg ctctcgttg 2040  
65 gtcagctga aactaggggc aaggttctat ccgttggcgc gcagcggcta catgtagggt 2100

catgagtcctt tcaccgtgga caaatccctt gaaaaaaaaa aaaggagtc c gggttaagcat 2160  
 tcattccggg tcaagtgtct ggttctgtga ataaactcta agattcaaga aaccttaatg 2220  
 aaagaaaacc ttgatgattc agagcaagga tgtggtcaca cctgtggctg gatctgtttc 2280  
 agccgcccc a gtgcatgggt agagtgggga gcagggtattg tttgttcaga ggtctcatct 2340  
 5 ggtatgtttc tgagggtgtt gccggctgaa tggtagacgt gtcgtttgtg tgtatgaggt 2400  
 tctgtgtctg tgtgtggctc ggtttgagtg tacgcatgtc cagcacatgc cctgccctgc 2460  
 tctcacctgc gtcttccgc cccag 2485  
  
 <210> 13  
 10 <211> 1984  
 <212> DNA  
 <213> Homo sapiens  
  
 <400> 13  
 15 gtgaggcctc ctcttcccca ggggggcttg ggtgggggtt gatttgcttt t gatgcattc 60  
 agtggttaata ttcttggtgc tctggagacc atgactgctc tgtcttgagg aaccagacaa 120  
 ggttgagacc ccttcttggt atgaagccgc acgggagggg ttgcacagcc tgaggactgc 180  
 gggctccacg caggctctgt ccagcgcca tgtccagagg cctcagggct cagcaggcgg 240  
 20 gagggccgct gccctgcacg atgagcatgt gaattcaaca ccgaggaagc acaccagctt 300  
 ctgtcacgtc acccaggttc cgttagggtc ctggggaga tggggctggt gcagcctgag 360  
 gccccacatc tcccagcagg cctcgacag gtggcctgga ctgggcgcct ctccagccca 420  
 ttgcccatcc cacttgcatg ggttctacac ccaaggacgc acacaccta atactgtgcc 480  
 aacctaatgt ggttcaactc agctggcttt tattgacagc agttactttt ttttttttaa 540  
 tactttaagt tctaggttac atgtgcacga cgtgcagggt agttacatat gtatacatgt 600  
 25 gccatgttgg tgtgtgcac ccattaaact atcatttaca ttaggtatat ctccataatgc 660  
 tateccctcc cactcccccc atcccatgac aggccttggt gtgtgatgtt cccaccctg 720  
 tgtccaaagt ttctcattgt tcagttccca cctgtgagtg agaacatgtg gtgtttggtt 780  
 ttctttcctt gcaatagttt gctcagagtg atggtttcca gcttcgtcca tgtccctaca 840  
 aaggacatga actcatcctt ttttatgact gcatagtatt ccgtgggtga tatgtgccac 900  
 30 attttcttaa tccagtctat catcgatgga catttgggtt ggttgcaagt ctttgcact 960  
 gtgaatagtg ccgcaataaa catacgtgtg catgtgtctt tatagcagca tgatttataa 1020  
 tcctttgggt atataccag taatgggatg gctgggtcaa atggtatttc tagttctaga 1080  
 tccttgagga atcaccacac tgtcttccac aatggttgaa ctagtttaca ctcccaccaa 1140  
 35 cagtgtaaaa gtgttctggt gctggagagg atgtggacag cagttatttt tttatgaaaa 1200  
 tagtatcact gaacaagcag acagttagt aaggatgctg caggaagcct gcaggccaca 1260  
 cagccatttc tctcgaagac tccgggtttt tctctgcat cttttgaaac tctagctcca 1320  
 attatagcat gtacagtgga tcaaggttct tcttcattaa ggttcaagt ctatattgaa 1380  
 ataatgttat gtaacagaaa caaaaatttc ttgtacacac aacttgctct gggatttggg 1440  
 40 ggaagtgtc ctgagctgg cggcacactg gtcagccctc tgggacagga tacctctggc 1500  
 ccattggtcat ggggcgctgg gcttgggcct gaggtgcaca cagtgcacca tgcccagctt 1560  
 cctgtggata ggatctgggt ctgggatcat gctgaggacc acagctgcca tgcctgtaaa 1620  
 gggcaccacg tggctcagag ggggcgaggt tcccagcccc agctttctta ccgtcttcag 1680  
 ttatttttcc ctaagagtct gagaagtggg gccgcgctg atggccttcg ttcgtcttca 1740  
 45 gctggcacag aattgcacaa gctgatggta aacactgagt acttataatg aatgaggaat 1800  
 tgcgtagca gtttaactgta gagagctcgt ctgttgga gaaatttaag ttttctattt 1860  
 aaccgctttg gagaatgtta ctttatttat ggctgtgtaa attgtttgac attcagtcct 1920  
 tcgtagacag atactacgta aaaagtgtaa agttaacctt gctgtgtatt tttccctatt 1980  
 ttag 1984  
  
 <210> 14  
 50 <211> 1871  
 <212> DNA  
 <213> Homo sapiens  
  
 <400> 14  
 55 gtgaggcccg tgccgtgtgt ctgtggggac ctccacagcc tgtgggcttt gcagttgagc 60  
 ccccggtgtc ctgccccctg caccgcagcg ttgtctctgc caagtccctc ctctctgccg 120  
 gtgctggatc cgcaagagca gaggcgcttg gccgtgcacc caggcctggg ggcgcagggg 180  
 60 caccttcggg agggagtggt taccgtgcag gccctgttcc tgcaagagc caccaggtt 240  
 acacacgtgg tgagtgcagg cgggtgacct gctcctgctg ctctttggaa agtcaagagt 300  
 ggcggctcct ggggccccag tgagaccccc aggagctgtg cacagggcct gcagggccga 360  
 ggcggcagcc tcctccccag ggtgcacctg agcctgcgga gagcaggagc tgcagagtga 420  
 gctggccac agcgttcgct gcggtcacgt tcctgcgtgg ggttgtttgg gatcggtggg 480  
 agaatttggg tttgtcagtg gctgtgtct tgaaccacgg agatggctag gagtgggttt 540  
 65 cagagttgat ttttgcgaat caaactaaaa tcaggcacag gggacctggc ctccagcacag 600

14 / 18

5 gggattgtcc aatgtgttcc ccctcaaggc cgccccacag agccggtggg cttgttttaa 660  
 agtgcgattt gacgagggac gagaacacct gaaagctgta aaggggaacc tcagaaaatg 720  
 tggccgccag ggggtgtttc aggtgctttg ctgggctgtg tttgtgaaaa cccattttga 780  
 cccgccctcc aagtccaccc tccaggtcca cctccaggc ccgccctggg ctgggggtat 840  
 gcctggcggt ccttgtgccc cagcccggag cacagcaggc tgtgcacatt taaatccact 900  
 aagattcact cggggggagc ccaggtccca agcaactgag ggctcaggag tcctgaggct 960  
 gctgagggga cagagcagac ggggaacgct gcttctgtgt ggcaagtccc tgaggggtgt 1020  
 ggccagggag gtggctcaga gtgtatgttg ggtccccc gggggcagaa ctctgtctct 1080  
 gatgagtcgg cagccatgta acaggaaggc gtggccacag ggagctggga atgcaccagg 1140  
 10 ggagctgcgc agctggccga ggtcccaggc ccaggccaca ggaagggcag ggggacgccc 1200  
 ggggccacag cagagccgcg aggaagggaa ggggatgccc aggccagagc agaggtctacc 1260  
 gggcacaggc gggctccctg agctgggtga gcgaggctca tgactcggcg agggaacctc 1320  
 cttgacgtga agctgacgac tgggtgttgc cagctcacag cccagccagg tccccgcgct 1380  
 gagcaggaac tcagaacctc cccctttgtc taaagcacag cagatgcctt cagggcatct 1440  
 15 agggaaaaac agggcaagtc gttgagaaac gtcttaaaag aaggtgggat ggtggcaatt 1500  
 tctgtccagc attttagctt gccccggacc acagatgagt ctataacggg attgtgtgtt 1560  
 tgccatgggg acacatgaga tggaccatca cagaggccac tggggctgca cctcccatct 1620  
 gagtccctggc tgtcccggtt ccaggccagg tcttgcctatg ctccctacc tgctctgccc 1680  
 gggagacagg gaaagcaccg cgaagtctgg agcagggtcg ggtccaggct cctcagagct 1740  
 20 cctgccaggc ccagcaccct gctccaaatc accacttctc tggggttttc caaagcattt 1800  
 aacaagggtg tcaggttacc tcctgggtga cggccccgca tcctggggct gacattgccc 1860  
 ctctgcctta g 1871

<210> 15  
 <211> 3801  
 <212> DNA  
 <213> Homo sapiens

30 <400> 15  
 gtgagcgcac ctggccggaa gtggagcctg tgcccggctg gggcaggtgc tgctgcaggc 60  
 ccgttgctgc caccctctgt tccgtgtggg gcaggcgact gccaatccca aagggtcaga 120  
 ggccacaggc tgcctctcgt cccatctggg gctgagcaga aatgcattct tctgtgggag 180  
 tgagggtgct cacaacggga gcagttttct gtgctatttt ggtaaaagga aatggtgcac 240  
 35 cagacctggg tgcactgagg tgtcttcaga aagcagctct gatccgaacc caagaccccc 300  
 gggccctgct gggcgtgagt ctctcaaac cgaacacagg ggcctctgct ggcatgagtc 360  
 cctctgaacc cgagaccctg gggccctgct gggcgtgagt ctctccgaac ccagagactt 420  
 cagggccctt ttggcgctga gtctctccgc tgtgagcccc acactccaag gctcatccac 480  
 agtctacagg atgccatgag ttcgatgaca cgtgtgacct atcaggggac agggccatgg 540  
 40 tgtggggggg gtctctacaa aattctgggg tcttgtttcc ccagagcccc agagctcaag 600  
 gccccgtctc aggtctcagc acaaatgaat tgaagatgga cacagatgca gaaatctgtg 660  
 ctgtttcttt tatgaataaa aagtatcaac attccaggca gggcaagggt gctcacacct 720  
 ataatccagc cactttggga ggcgaggtg ggtggatcac ttgaggccag gagtttgagg 780  
 ccaacctaac caacatagtg aaattccatt tctacttaaa aaatacaaaa attagccttg 840  
 cctggtggca cagcctgtga gtccccgcta tgccggaggc tgaggcagga gaatcatttg 900  
 45 aaccaggag gcagaggttg cagtgagccg agatcacacc actgcactcc agcctgggca 960  
 acagagttag acttcatctt aaaaaaaaa aaaaaagtat cagattcca aaacctagat 1020  
 ggacaggtgt ttttttattc tgccttcga taatatttac tgggtgctgtg ctgagggccg 1080  
 gaaactggggg tgcttccctc tgaaaggcac accttcattg gaagagaaat aagtgggtga 1140  
 tggttgttaa accagaggtt taaactgggg tctgtctgtt ctgagttaac agtccagatc 1200  
 50 tggactttgc ctctttccag aatgctccct ggggtttgtt tcatggggga gcagcagggt 1260  
 tggacaccct cgtgatgggg gagcagcagg tgcagacgcc ctcatgatgg gggagtggca 1320  
 ggtgcagaca cctttgtgca tgggtcccag catgtccctg ttgcagctcc ctcccccaa 1380  
 ggatgccggt ctctgtgtct cccacagtc cctgcttccc tctcacagcc ttaacctggtc 1440  
 ctggcctcca ctggctttgt ctgcatgatt tccacatttc ctgggctccc agcactctt 1500  
 55 cgctctccc aggcacctct gcagtgtctg ccataccagt cagctgtgaa ctgtccactg 1560  
 cttattttgc tccccatgaa atgtattttt taggacaggc acccctggtt ccagcctctg 1620  
 gcacagcatc agtgaatggt attgaaggac aaaggacaga caaacaatc agggaaatgg 1680  
 gttctctcta aacacattgc aaagccacag aggtagtgtc aggtatgggtg ggcacagggt 1740  
 60 catcagatgt ggggtccaatg ccagaatatt ctgtgtctcc aaaggccact tggtcagagt 1800  
 gtgtgcttgc agaggtggct ctaaaagctc agcagtggag gcagtgggtc gccatactca 1860  
 ggggtgaactc acatcctctg tgtctgaagt atacagcaga ggcttgaagg gcattctggga 1920  
 gaagaaaaa ggcaaatga ttaagaaaag tgaagaaagg aaagtggtaa gatgggaatt 1980  
 ttcttgtcca gatttttagt tccccaaaca cagctcagat ggtagaatgt ggtcagaact 2040  
 gatggacaga acaatagaac aaaacggaag ccctatctct cagaacgtg tgttaattgt 2100  
 65 gtatgtggca cagctgatgt aaaagagagt gtgtgtgtaa ttttttttcc tgagaaaaat 2160

15 / 18

5 gactggaagc aaataagttg tgtctttaca gcatatacca gagcagattc taggtagaag 2220  
 aggagacaca tgcaaaacac accagcaaca gaaataaac aaaagactca aaggggaaggg 2280  
 aggtgaacgt tccctgggtt ggtgttgggg aaggacacac aggggagcggt atgaaaccag 2340  
 tgaggcaacg ggcattgctt tcactgcaga gaaactcagc ttgcttgagc cacagtgaag 2400  
 atggccattc cctggagcgt ttgtgcacgt gatttattta aggcgccttg tgaggtcctg 2460  
 cacattcatc ctctcacttt gttctcctaa ccacctgaga ggtagaggag gaaaggctcc 2520  
 aggggagcag ccgccccttg tcaccagcgt ggcaaaaggg atgcatgatt gcagcctggc 2580  
 ctcttgctcc ggggcccttg ctctgccga ggacccaca caagtcagac ccataggctc 2640  
 10 aggggtgagc ggagcccaag gtcgtgttgg ggatggctgt gaaagaagaa atggagctct 2700  
 gatgcacact tgggaaggct ctaccagcag cgtcaaaagaa atgcatgtga aactgacagc 2760  
 gagaccatc cctcaaagaa acgcacgtga aactgatggc gagacctgtc cccatccctc 2820  
 atgctggctc cttttctggg cttgccaaaga gccagcatca ggttgaggca agctggaaaag 2880  
 acttttctgg aaagcagctt gtttgcatgg aagtcctcac aatgtcctgt gtcttccag 2940  
 taattccact tctgaagtga ccagacatta tcacgggtct tatttaccat ttccagtggt 3000  
 15 ccaggcaggg ggacttgcca cagcaagtca cgaacctgcc caaatacagg gctaaggaga 3060  
 tattatgcat cacaaaactt gctctgccat taacatctt tcaaagaatt ttgaaagaat 3120  
 gtttaatggc acaaacctgt tatttcaatg tagcagtggt caaagctgga tgaataagaa 3180  
 cacacccagc gagcctgccg tgaatgtcat gtgtgttcat ctttggaatc ggacatacat 3240  
 20 gggcagtgag tgggtggtag gccctggagg acatcggttg gatgcctcca tctgcccct 3300  
 ctggagacac catgtgtgcc acgtgcactc actggagccc tgtttagctg gtgccacctg 3360  
 gctcttccat ccctgagatt caaacacagt gagattcccc acgcccacat cagtgtcttc 3420  
 ccacaaaaaa cctgagtcac acctgtgttc actcgaggga cgccggggag ccagggtctc 3480  
 acagtttatt atgtgttttt ggctgagtta tgtgcagatc tcatcagggc agatgatgag 3540  
 25 tgcacaaaca cggccgtgag aggtttggat acactcaaca tcaactagcca ggtccctggt 3600  
 gagtttggtc atgcagatgc tggatggcat gtagcatttg gtagtccatgg agtgagcacc 3660  
 cagcccccctc gggctgcagc catgcccaca gccaggacaa ggaagcgagg ggaaggcagg 3720  
 aggtctcttg gagcaagctt tgcaggaggg ggctgggtgt ggggcaggga cctgtgtctg 3780  
 acattccccc ctgtgtctca g 3801

30 <210> 16  
 <211> 880  
 <212> DNA  
 <213> Homo sapiens

35 <400> 16  
 gtgagcaggc tgatgggtcag cacagagttc agagtccagg aggtgtgtgc gcaagtatgt 60  
 gtgtgtgtgt gtgcgcgcgt gcctgcaagg ctgatggtga ctggctgcac gtaagagtgc 120  
 acatgtacgc atatacacgt gagcacatac atgtgtgcat gtgtgtacat gaaggcatgg 180  
 40 cagtgtgtgc acaggtgtgc aagggcacaa gtgtgtgcac atgcgaatgc acacctgaca 240  
 tgcatgtgtg ttctgtgcaca gtcgtgtggg cattcacgtg aggtgcatgc gtgtgggtgt 300  
 gcagtgtgag tagcatgtgt gcacataaca tgtattgagg ggtcctcgtg ttcaccccgc 360  
 taggtcctca gcaccagtgc cactccttac aggatgagac ggggtccagc gccttggttg 420  
 gctgagggtc tgaagctgca gccctgaggg cattgtccca tctgggcacg cgcgtccact 480  
 45 cctctcctg tgggcttctg tgtccactcc cctctcctg tgggcattta catccactcc 540  
 actccctctc tctgtgggc atccgcgtcc actccccctc tctgtgggca tctgcgtcca 600  
 cctccctctc ctgtgggcat ttgcgtccac tccctctcct ggttctctcc tgtcttggtc 660  
 gagcctcggg ggcaggcaga tgacacagag tcttgactcg ccagggtggg ttcgcagctg 720  
 cgggttgagg gccaggccgg atttcactgg gaagagggat agtttcttgt caaaatgttc 780  
 50 ctcttctctg ttccatctga atggatgata aagcaaaaag taaaaactta aaatcccaga 840  
 gaggtttcta ccgtttctca ctcttctctg gcgactctag 880

55 <210> 17  
 <211> 3186  
 <212> DNA  
 <213> Homo sapiens

60 <400> 17  
 gtgagccgcc accaaagggg gcaggccag cctccaggga ccctccgcgc tctgctcacc 60  
 tctgaccggg ggtctcacct tggaaactct gggttttagg ggcaaggaaat gtcttacgtt 120  
 ttcagtgtgt ctgctgcctg tgcacagttc tgttcgcgtg gctctgtgca aagcacctgt 180  
 tctccatctc tgggtagtgg taggagccgg tgtggcccca ggtgtcccca ctgtgcctgt 240  
 gctatggcgt tgggacgtca tggaggccat ccaggggcag caggggcatg gggtaaaagag 300  
 atgtttatgg ggagtcttag cagaggaggc tgggaagggt tctgaacagt agatgggaga 360  
 65 tcagatgccc gaggagattg ggggtctcag aaagaggggc gaggtgggtg caggtgaggg 420  
 tcgctggccc caccgccggg aaggtgcagc agagctgtgg ctccccacac agcccggcca 480

	gcacctgtgc	tctgggcatg	gctgtgctcc	tggaaacgttc	cctgtcctgg	ctggtcaggg	540
	gggtgcccctg	ccaagaatcg	acaactttat	cacagaggga	agggccaatc	tgtggaggcc	600
	acagggccag	cttctgectg	gagtcagggc	aggtgggtggc	acaagcctcg	gggctgtacc	660
5	aaagggcagt	cgggcaccac	aggcccgggc	ctccacctca	acaggcctcc	cgagccactg	720
	ggagctgaat	gccaggaggc	cgaagccctc	gccccatgag	ggctgagaag	gagtgtaggc	780
	atctgtgtta	cccagggccg	aggctgcgcg	aattaccctg	cacacttgat	gtgaaatgag	840
	gtcgtcgtct	atcgtggaaa	cccagcaagg	gctcacggga	gagttttcca	ttacaaggtc	900
	gtaccatgaa	aatgggtttt	aacccgagtg	cttgcgcttc	catgctctgg	cagggagggc	960
10	agagccacag	ctgcatgtta	ccgcctttgc	accagctcca	gaggcttggg	accaggctgt	1020
	ctcagttcca	gggtgcgtcc	ggctcagacc	gcccctctct	ctgccttctc	tctctgcctc	1080
	aaatcttccc	tcgtttgcat	ctccctgacg	cgtgcctggg	ccctcgtgca	agctgcttga	1140
	ctcctttccg	gaaacccttg	gggtgtgctg	gatacaggtg	ccactgagga	ctggaggtgt	1200
	ctgacactgt	ggttgacccc	agggtccacg	tggcgtgctt	ggggcctcct	tgggccatga	1260
	tgaggtcaga	ggagttttcc	caggtgaaaa	ctcctgggaa	actcccaggg	ccatgtgacc	1320
15	tgccacctgc	tcctcccata	ttcagctcag	tctgtctctc	atttcccac	caggggtctct	1380
	agctccgagg	agctcccgtg	gagggcctgg	gctcagggca	gggcggctga	gtttcccac	1440
	ccatgtgggg	acccttgggt	agtgccttga	tgggttagcc	ctgaggaggc	cgagatgcga	1500
	tggggccacg	gccgtttcca	aacacagagt	caggcacgtg	gaaggccag	gaatcccctt	1560
20	ccctcgaggc	aggagtggga	gaacggagag	ctgggcccgc	atttcacggc	agccaggtct	1620
	cagtgggcga	ggctgtggtg	gtccacgtgg	cgctgggggc	gggtctgat	tcaaaccgc	1680
	tggggctcgg	ccttctctgg	ccgtgtctgg	cgccctcca	cacgggcttg	gggtggacgc	1740
	ccgacctct	agcaggtggc	tattctctcc	tttgggaagag	agcccctcac	ccatgctagg	1800
	tgtttccctc	ctgggtcagg	agcgtggccg	tgtggcaacc	ccgggacctt	aggcttatct	1860
25	atctgtttaa	aaacattctg	ggcctggctt	ccgttgttgc	taaatgggga	aaagacatcc	1920
	cacctcagca	gagttactga	gaggctgaaa	ccgggtgctt	ggcttgactg	gtgtgatctc	1980
	aggtcattcc	agaagtggct	caggaagtca	gtgagaccag	gtacatgggg	ggctcaggca	2040
	gtgggtgaga	tgaggtacac	ggggggctca	ggcagtgggt	gaggccaggt	acatgggggg	2100
	ctcaggcact	gggtgagatg	aggtacacgg	ggggctcagg	cagaggttca	gaccaggtac	2160
30	acgggggctc	tgatcacacg	cacatatgag	cacatgtgca	catgtgctgt	ttcatggtag	2220
	ccaggtctgt	gcacacctgc	cccaaagtcc	caggaagctg	agaggccaaa	gatggaggct	2280
	gacagggtcg	gcgcggtggc	tcacacctgt	agtcccagca	ctttgggagg	ccgaggcgag	2340
	aggatcccctt	gagcccaggga	gtttaagacc	agcctgagca	acatagtaga	accccatctc	2400
	tatgaaaaat	aaaaacaaaa	attagctgaa	catggtggtg	tgccctgta	gttccaatcc	2460
35	ttgggaggct	gaagtgggag	gatcacttga	gcccaggagg	tgggaagctg	agtgagctga	2520
	gattgcacca	ctgtactgca	gcctgggtga	cagagtgaga	gccccatcca	acaacaacaa	2580
	agaagactga	caaatgcagt	ctcttgaaa	gaaacattta	gtaggaactt	aacctacaca	2640
	cagaagccaa	gtcgtgtgtc	cgggtcagt	gagatgagat	gatgggtcct	cacaccatca	2700
	ccccagacct	agggtttatg	caccacaggg	gcgggtggct	cagaagggat	gcgcaggacg	2760
40	ttgatatacg	atgacatcaa	ggttgtctga	cgaaggggag	gattcatgat	aagtacctgc	2820
	tggtacacaa	ggaacaaatg	ataaactgga	aaccttagag	gccttcccgc	aacaggggct	2880
	aatcagaagc	cagcatgggg	ggctggcatc	caggatggag	ctgcttcacg	ctccacatgc	2940
	gtgttcatac	agatggtgca	cagaaacgca	gtgtacctgt	gcacacacag	acacgcagct	3000
	actgcacac	acaagcacac	acacagacat	gcattgcattg	atccgtgtgt	gtgcacctgt	3060
45	gcccattgag	aaacccatgc	atgtgcattc	atgcacgcac	acagggcaccg	gtggggcccat	3120
	gcccacaccc	acgagcaccg	tctgattagg	aggcctttcc	tctgacgctg	tccgccatcc	3180
	tctcag						3186
	<210> 18						
	<211> 781						
50	<212> DNA						
	<213> Homo sapiens						
	<400> 18						
55	gtatgtgcag	gtgcctggcc	tcagtggcag	cagtgccctg	ctgctgggtg	tagtgtgtca	60
	ggagactgag	tgaatctggg	cttaggaagt	tcttaccctt	tttcgcacat	ggaagtgggt	120
	taacccaacc	actgtcaggc	tcgtctgccc	gccctctcgt	gggtgtgagca	gagcacctga	180
	tggaggggac	aggagctgtc	tgggagctgc	catecttccc	accttgcctc	gcctggggaa	240
	gcgctggggg	gcctgtcttc	tcctgtttgc	cccatgggtg	gatttggggg	gcctggcctc	300
60	tcctgtttgc	cctgtgggtg	gattgggctg	tctcccgctc	atggcactta	gggcccctgt	360
	gcaaaccagg	gccaagggct	tagggaggag	ccaggcccag	gctaccccac	ccctctcagg	420
	agcagaggcc	gcgtatcacc	acgacagagc	cccgcccggt	cctctgcttc	ccagtcaccg	480
	tcctctgccc	ctggacactt	tgtccagcat	cagggagggt	tctgatccgt	ctgaaattca	540
	agccatgtcg	aacctgcggg	cctgagctta	acagcttcta	ctttctgttc	tttctgtgtt	600

gtggaaattt cacctggaga agccgaagaa aacatttctg tcgtgactcc tgcggtgctt 660  
 gggtcgggac agccagagat ggagccaccc cgcagaccgt cgggtgtggg cagctttccg 720  
 gtgtctcctg ggaggggagc tgggctgggc ctgtgactcc tcagcctctg ttttcccca 780  
 g 781

<210> 19  
 <211> 536  
 <212> DNA  
 <213> Homo sapiens

<400> 19  
 gcaagtgtgg gtggaggcca gtgcgggccc cacctgccca ggggtcatcc ttgaacgccc 60  
 tgtgtggggc gagcagcctc agatgtctgt gaagtgcaga cggcccccgg cctgaccctg 120  
 ggggcctgga gccacgcttg cagccctatg tgattaaacg ctggtgtccc caggccacgg 180  
 agcctggcag ggtccccaac ttcttgaacc cctgcttccc atctcagggg cgatggctcc 240  
 ccacgcttgg gagccttctg acccctgacc tgtgtcctct cacagcctct tccctggctg 300  
 ctgccctgag ctctctgggt cctgagcaag ttctctcccc gccccggcgc tccagcgtca 360  
 ccgggctgcc tgtctgctcg ccccggtgga ggggtgtctg tcccttccact gaggttccca 420  
 ccagccaggg ccacgaggtg caggccctgc ctgcccggcc acccacacgt cctaggaggg 480  
 ttggaggatg ccacctctgg cctcttctgg aacggagtct gatttttgcc ccgcag 536

<210> 20  
 <211> 3179  
 <212> DNA  
 <213> Homo sapiens

<400> 20  
 atctcatgtt tgaatcctaa tgtgcactgc atagacacca ctgtatgcaa ttacagaagc 60  
 ctgtgagtga acggggtggg gtgcagtgcg ggcccatggc ctggctgtgc atttacggaa 120  
 gtctatgagt gaatggggtt gtggtcagtg cgggcccacg gcctggctgg gcctggagg 180  
 ttctctgatg tgtgagggag gagggggaag agggtagggg atagacagtg ggagcccca 240  
 ccttggaaga cataacagta agtccaggcc cgaaggcgag cagggatgct gggggcccg 300  
 ctggggcggc ggggatgatg gagggcctgg ccagggtggc agggatgatg gggggcccg 360  
 ctgggggtggc aggggtgatg gggggggctg gtctgggtgg cggggaagat ggggaagcct 420  
 ggctgggccc cctcctcccc tgcctccccc ctgcagccgt ggatccggat gtgttctcct 480  
 ggtgcacatc ctctgggcca tcagctttca tggaggtggg gggcaggggc atgacaccat 540  
 cctgtataaa atccaggatt cctcctctg aacgcccaca ctcaggttga aagtcacatt 600  
 ccgctctctg ccattctctt aagagtagac caggattctg atctctgaag ggtgggtagg 660  
 gtggggcagt ggagggtgtg gacacaggag gcttcagggt ggggctgggt atgtctcttc 720  
 atcctcttat catctcccag tctcatctct catcctctta tcctctccca gtctcatctg 780  
 tcttctctct atctcccagt ctcatctgtc atcctcttac catctcccag tctcatctct 840  
 tctctctta tctcttagtc tcattccagc ttacctccca gggcgggtgc caggctcgca 900  
 gtggagctgg acatacgtcc ttctcaggc agaaggaaact ggaaggattg cagagaacag 960  
 gaggggcggc tcagagggac gcagcttgg ggtgaagaaa cagccctcc tcagaagtgt 1020  
 gcttgggcca cagaaaccg agggccctgc gtgagtggct ccagagcctt ccagcaggtc 1080  
 cctgggtggg ccttatggta tggcgggtgc ctactgagtg caccttgagc agggcttctg 1140  
 gtttgagtgc agcccggagc tgcctgggtg cggggtgggg gcttatggcc actggatatg 1200  
 gcgtcattta ttgctgctgc ttccagagaat gtctgagtga ccgagcctaa tgtgtatgg 1260  
 gggcccaagt ccacagactg tgtcgtaaat gcactctggt gcctggagcc cccgtatagg 1320  
 agctgtgagg aaggaggggc tcttggcagc cggcctgggg gcgcctttgc cctgcaaaact 1380  
 ggaaggggagc gggcccgggc gccgtgggag gacgacctca agtgagaggt tggacagaac 1440  
 agggcgggga cttcccagga gcagaggccg ctgctcaggc acacctgggt ttgaatcaca 1500  
 gaccaacagg tcaggccatt gttcagctat ccattctcta caaagctcca gattcctgtt 1560  
 tctccgggtg ttttttgggt aaattttact caggattact tatatttttt gctaaagtat 1620  
 tagaccctta aaaaagggtat ttgctttgat atggcttaac tactaagca cctactttat 1680  
 ttgtctgttt ttatttatta ttattattat tattagagat ggtgtctact ctgtcaccca 1740  
 ggttgttagt gcagtggcac agtcatggct cgctgtagcc gcaaaccctc aggcctcaagt 1800  
 gatectccgg cctcagcttc ccagagtgtc gggattacag gtgtgagcca ctgccccttg 1860  
 ctggcacttt taaaaaccac tatgtaaggt caggctccagt ggcttccaca cctgtcatcc 1920  
 cagtgtttt ggaagccgag gcagaaggat tgtctgaggc caggagtgtt agaccagcat 1980  
 gggtaacata gggagacccc atctctacaa aaaaatgcaa aagtattccg ggcgtgggg 2040  
 ccagcatctg tagtcccagc tgcctgggag gctgagtggg aggatcgctt gagccccgga 2100  
 ggtcatggct gcagttagct gtgattgtac catcgactc cagcctgggc aacagagtga 2160  
 gacctgtctt caaaaaaaaa aaaaaaaaaa gaaggagaag gagaagagaa gaagaaggaa 2220  
 gaaggaaaaga gaagaagaag gaagaaggaa gaaagaagga gaaggaggcc tgcagggtgc 2280

18 / 18

taggtagact gtcaaatttc agagcaaaat gaaaataaca aagttttaaa gggaaagaaa 2340  
aaccccagct ctttggactt ccttaggcct gaacttcac tcaagcagct tccttccaca 2400  
gacaagcgtg tatggagcga gtgagttcaa agcagaaagg gaggagaagc aggcaagggc 2460  
ggaggctgtg ggtgacacca gccaggaccc ctgaaaggga gtggttggtt tcctgcctca 2520  
5 gccccacgct cctgccggtc ctgcacctgc tgtaaccgtc gatgttggtg ccagggtgccc 2580  
acctgggaag gatgctgtgc agggggcctt ccaaaactttg gtgggtttca gaagcccag 2640  
gcacttggtg caggcacaaat tacagccctt ccccaaagat gcccacgtcc ttctcctgga 2700  
acctgtgaat gtgtcaccg caaggcagag gctggtgaag gctgcaggtg gaatcacggc 2760  
tgccagtcag ccgatcttaa ggtcatcctg gattatctgg tgggcctgat atggccacaa 2820  
10 gggtccttag aagtgaagaga gggaggcagg ggagagtcag agaggggacg tgagaaggac 2880  
cactggccac tgctggcttt gagatggagg agggggctcc cagccaagga atgggggcag 2940  
ccgtccatg ctggaaaagc aagcaatcct ccccggtcct gagggcacac ggccctgccc 3000  
acgcctcgat ttcaggccag tgggacctgt ttcagctttc cggcctccag agctgtaaga 3060  
tgatgcgtt gtgttcagcc actaagctgc agtgattcgt cacagcagca aatggaatag 3120  
15 cagtacaggg aaatgaatac agggacagtt ctccagtgta ctctcagccc acccctggg 3179